THE BEST ESTIMATE OF A MULTIVARIATE NORMAL DISTRIBUTION

By

HONG-I YANG

Bachelor of Arts
National Taiwan University
Taipei, Taiwan, Republic of China
1956

Bachelor of Arts
Montana State University
Missoula, Montana
1961

Master of Science
Iowa State University
Ames, Iowa
1964

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1974

THE BEST ESTIMATE OF A MULTIVARIATE

NORMAL DISTRIBUTION

Thesis Approved:

J. Leroy Folks

Thesis Adviser

K. S. Mansky

DR. W. D. WARDE WDW

Y. C. Lu

N. N. Durham

Dean of the Graduate College

ACKNOWLEDGMENTS

I am deeply indebted to Professor J. Leroy Folks for suggesting this problem and his guidance in the preparation of this thesis.

An expression of gratitude is extended to Professors Lyle D. Broemeling, Ignancy I. Kotlarski, Yao C. Lu, and William Warde for serving on my advisory committee.

I would like to extend my sincere appreciation to the entire faculty as well as my fellow students for the friendship and help they have given me.

Finally, I would like to express my thanks to Mrs. Joyce Gazaway for her professional typing of a very difficult thesis.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1.1 Background and Scope

The best (minimum variance unbiased) estimate of the probability F(h), where

$$F(h) = \int_{-\infty}^{h} \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{x - \mu}{\sigma}\right)^2} dx, \ \mu \varepsilon R, \ \sigma > o; \ x \varepsilon R, \tag{1.1}$$

has been investigated by Folks, Pierce and Stewart [1] for (i) known mean and unknown variance, (ii) unknown mean and known variance, and (iii) unknown mean and unknown variance. In that paper the estimates of the probabilities of Equation (1.1) for various cases of the parameters being expressed in terms of either the cumulative distribution of the standard normal variate or the cumulative Student's distribution are hence made more useful to the non-mathematically oriented reader. An attempt is made in the same vein by this writer to extend the integrand of Equation (1.1) to a bivariate normal density. That is, if necessary and possible, the expectation of the conditional density, if it exists, has resulted from the search of the best estimates of the probabilities of a bivariate normal distribution for various cases of the parameters are to be transformed to some recognizable distributions of which tables are available. Specifically, this writer has attempted to find the best estimates of F(h,k), where

$$F(h,k) = \int_{-\infty}^{h}\int_{-\infty}^{k} \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp -\left\{\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\} dydx \tag{1.2}$$

$\mu_x, \mu_y \in R, \sigma_x, \sigma_y > o, o < |\rho| < 1, x, y \in R,$

for the following cases.

	ρ	μ_x	μ_y	σ_x	σ_y
1	u	u	u	u	u
2	k	u	u	k	k
3	o	u	u	k	k
4	o	k	k	u	u
5	o	u	k	u	k
6	o	u	u	u	u

(u = unknown, k = known)

Small sample problems are investigated and numerical examples are given for each case. An attempt is made to find the best estimate of

$F(h_1,\ldots,h_p)$

$$= \int_{-\infty}^{h_1} \cdots \int_{-\infty}^{h_p} \frac{1}{(2\pi)^{p/2}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(X-\mu)'\Sigma^{-1}(X-\mu)\right] dx_p \ldots dx_1, \tag{1.3}$$

$\mu \in R^p$, Σ positive definite,

for unknown μ and unknown Σ.

1.2 Method of Estimation

The basic method used throughout this paper to derive the best estimate of the probability $F(h,k)$ as given by Equation (1.2) for various cases of the parameters is the application of the Rao-Blackwell and Lehmann-Scheffé theorems [2] and [3]. Namely if T is a complete sufficient statistic for θ, where θ is a d-dimensional parameter vector ($d \geq 1$), and if $\tilde{g}(\theta)$ is an unbiased estimator of $g(\theta)$ then the minimum variance unbiased estimator $\hat{g}(\theta)$ of $g(\theta)$ is given by

$$\hat{g}(\theta) = E[\tilde{g}(\theta)|T].$$

A more detailed description of the method as generalized by Basu [4] is given below.

For our case, let $\{(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)\}$ be a random sample of size n from a given bivariate normal distribution. Partition this sample into two subsamples $\{(X_1,Y_1)\}$ and $\{(X_2,Y_2), \ldots, (X_n,Y_n)\}$. Let

$$\begin{aligned} \tilde{g}(\theta) &= 1 \quad \text{if } X_1 < h \text{ and } Y_1 < k, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Then $E[\tilde{g}(\theta)] = F(h,k)$. Let T be a complete sufficient statistic for θ, if one exists, from the entire sample. By Rao-Blackwell and Lehmann-Sheffé theorems the best estimate of F(h,k) is given by $E[\tilde{g}(\theta)|T]$. But

$$E[\tilde{g}(\theta)|T] = P(X_1 \leq h,\ Y_1 \leq k|T).$$

One approach to evaluate this probability is to find the conditional density $f_{X_1,Y_1|T}(x_1,y_1|t)$ if it exists. Let T* be a complete sufficient statistic for θ from the subsample $\{(X_2,Y_2), \ldots, (X_n,Y_n)\}$. One can find the joint probability density function (p.d.f.) $f_{X_1,Y_1,T^*}(x_1,y_1,t^*)$ and hence the joint p.d.f. $f_{X_1,Y_1,T}(x_1,y_1,t)$ from which the conditional density $f_{X_1,Y_1|T}(x_1,y_1|t)$ can be obtained. Then

$$\begin{aligned} E\ \tilde{g}(\theta)|T &= P(X_1 \leq h_1,\ \ Y_1 \leq k|T) \\ &= \int_{-\infty}^{h}\int_{-\infty}^{k} \frac{f_{X_1,Y_1,T}(x_1,y_1,t)}{f_T(t)}\, dy_1 dx_1 \end{aligned} \tag{1.4}$$

is the desired estimate provided the integral exists. Now, if necessary and possible, Equation (1.4) is to be transformed to some special distribution so that the desired probability can be obtained from existing tables.

CHAPTER II

EXTENSION TO BIVARIATE NORMAL DISTRIBUTION AND SPECIAL CASES

It has been shown in Folks, Pierce and Stewart [1] that the best estimate $\hat{g}(\theta)$ of Equation (1.1) in a sample of size n is

(i) μ unknown, σ^2 known

$$\hat{g}(\theta) = F_z(\sqrt{n}(h-\bar{x})/\sigma\sqrt{n-1})$$

where F_z is the cumulative distribution of the standard normal variate.

(ii) μ known, σ^2 unknown

$$\begin{aligned} \hat{g}(\theta) &= 0, \quad h < \mu - \sqrt{n}\hat{\sigma} \\ &\ 1, \quad h < \mu + \sqrt{n}\hat{\sigma} \\ &= F_{t,n-1}\left\{(n-1)^{\frac{1}{2}}(h-\mu)/\left[n\hat{\sigma}^2-(h-\mu)^2\right]^{\frac{1}{2}}\right\} \quad \text{otherwise} \end{aligned}$$

where $F_{t,n-1}$ denotes the cumulative Student's t distribution with n - 1 degrees of freedom and where

$$n\hat{\sigma}^2 = \sum_{i=1}^{n} (X_i-\mu)^2.$$

(iii) μ and σ^2 unknown

$$\hat{g}(\theta) = 0, \quad h < \overline{x} - (n-1)S/\sqrt{n}$$

$$1, \quad h > x + (n-1)S/\sqrt{n}$$

$$= F_{t,n-2}\left\{\left[n(n-2)\right]^{\frac{1}{2}}(h-\overline{x}) / \left[(n-1)^2S^2 - n(h-\overline{x})^2\right]^{\frac{1}{2}}\right\} \quad \text{otherwise}$$

where

$$(n-1)S^2 = \sum_{i=1}^{n} (X_i - X)^2.$$

Extension is developed in this chapter to find the best estimate of F(h,k) as given by Equation (1.2) for various cases of the parameters. Also, small sample problems are considered and examples are given for each case.

2.1 ρ, μ_x, μ_y, σ_x^2, σ_y^2 Unknown

Let $\{(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)\}$ be a random sample of size n from a population with density

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2(1-\pi^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\} \qquad (2.1.1)$$

$$\mu_x, \mu_y \in R, \sigma_x, \sigma_y > 0, \ 0 < |\rho| < 1, \ x, y \in R.$$

Let

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad \overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i,$$

$$A_{11} = \sum_{i=1}^{n} (X_i - \overline{X})^2, \qquad A_{22} = \sum_{i=1}^{n} (Y_i - \overline{Y})^2$$

$$A_{12} = \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})$$

$T = \{\overline{X}, \overline{Y}, A_{11}, A_{12}, A_{22}\}$ is complete sufficient for μ_x, μ_y, σ_x^2, σ_y^2, and ρ (A.1, Appendix). The p.d.f. of T is given by

$$f_T(\overline{x}, \overline{y}, a_{11}, a_{12}, a_{22})$$

$$= \frac{n}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{n}{2(1-\rho^2)}\left[\left(\frac{\overline{x}-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{\overline{x}-\mu_x}{\sigma_x}\right)\left(\frac{\overline{y}-\mu_y}{\sigma_y}\right) + \left(\frac{\overline{y}-\mu_y}{\sigma_y}\right)^2\right]\right\} \cdot \tag{2.1.3}$$

$$\frac{\left(a_{11}a_{22}-a_{12}^2\right)^{\frac{1}{2}(n-4)} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{a_{11}}{\sigma_x^2} - 2\rho\frac{a_{12}}{\sigma_x\sigma_y} + \frac{a_{22}}{\sigma_y^2}\right)\right]}{2^{n-1}\sigma_x^{n-1}\sigma_y^{n-1}(1-\rho^2)^{\frac{1}{2}(n-1)} \sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}$$

$$\overline{x}, \overline{y}, a_{12} \in R, \; a_{11} \geq 0, \; a_{22} \geq 0,$$

where the second factor to the right of the equality sign is the Wishart density [5].

Let

$$\tilde{g}(\theta) = 1 \quad \text{if } X_1 < h \text{ and } Y_1 < k,$$

$$= 0 \quad \text{otherwise}$$

where $(h,k) \in R^2$ is fixed. $\tilde{g}(\theta)$ is unbiased for $F(h,k)$. The best estimator $\hat{g}(\theta)$ of $F(h,k)$ is then given by

$$\hat{g}(\theta) = E[\tilde{g}(\theta)|T] = P(X_1 < h, Y_1 < k|T).$$

In order to evaluate this probability, partition the random sample $\{(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)\}$ into two subsamples, $\{(X_1,Y_1)\}$ and $\{(X_2,Y_2), \ldots, (X_n,Y_n)\}$. Let

$$\overline{X}^* = \frac{1}{n-1}\sum_{i=2}^{n} X_i, \quad \overline{Y}^* = \frac{1}{n-1}\sum_{i=2}^{n} Y_i$$

$$A_{11}^* = \sum_{i=2}^{n}(X_i-\overline{X}^*)^2, \quad A_{22}^* = \sum_{i=2}^{n}(Y_i-\overline{Y}^*)^2,$$

$$A_{12}^* = \sum_{i=2}^{n}(X_i-\overline{X}^*)(Y_i-\overline{Y}^*).$$

$T^* = \{\overline{X}^*, \overline{Y}^*, A_{11}^*, A_{12}^*, A_{22}^*\}$ is complete sufficient for μ_x, μ_y, σ_x^2, σ_y^2, and ρ. Since (X_1,Y_1), $(\overline{X}^*,\overline{Y}^*)$, and $(A_{11}^*, A_{12}^*, A_{22}^*)$ are stochastically independent their joint p.d.f. is given by

$$f_{X_1,Y_1,T^*}(x_1,y_1,\overline{x}^*,\overline{y}^*, a_{11},a_{12},a_{22})$$

$$= \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x_1-\mu_x}{\sigma_x}\right)\left(\frac{y_1-\mu_y}{\sigma_y}\right) + \left(\frac{y_1-\mu_y}{\sigma_y}\right)^2\right]\right\}\cdot$$

$$\frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{n-1}{2(1-\rho^2)}\left[\left(\frac{\bar{x}^*-\mu_x}{\sigma_x}\right)^2 - 2\left(\frac{\bar{x}^*-\mu_x}{\sigma_x}\right)\left(\frac{\bar{y}^*-\mu_y}{\sigma_y}\right) + \left(\frac{\bar{y}^*-\mu_y}{\sigma_y}\right)^2\right]\right\}\cdot \tag{2.1.5}$$

$$\frac{\left(a^*_{11}a^*_{22}-a^{*2}_{12}\right)^{\frac{1}{2}(n-5)} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{a^*_{11}}{\sigma_x^2} - 2\,\frac{a^*_{12}}{\sigma_x\sigma_y} + \frac{a^*_{22}}{\sigma_y^2}\right)\right]}{2^{n-2}\sigma_x^{n-2}\sigma_y^{n-2}(1-\rho^2)^{\frac{1}{2}(n-2)}\sqrt{\pi}\,\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n-3}{2}\right)}\,.$$

Equation (2.1.5) is to be transformed to the $X_1Y_1\bar{X}\,\bar{Y}A_{11}A_{12}A_{22}$ - space.

Let (A.2, Appendix)

$$X_1 = X_1, \quad Y_1 = Y_1, \quad \bar{X}^* = \frac{n\bar{X}-X_1}{n-1}, \quad \bar{Y}^* = \frac{n\bar{Y}-Y_1}{n-1}$$

$$A^*_{11} = A_{11} - \frac{n}{n-1}(\bar{X}-X_1)^2, \quad A^*_{22} = A_{22} - \frac{n}{n-1}(\bar{Y}-Y_1)^2, \tag{2.1.6}$$

$$A^*_{12} = A_{12} - \frac{n}{n-1}(\bar{X}-X_1)(\bar{Y}-Y_1).$$

Since $a^*_{11}\,a^*_{22} - a^{*2}_{12} > 0$, Equations (2.1.6) give

$$1 - \frac{n}{n-1}\left[\frac{a_{22}(x_1-\bar{x})^2 - 2a_{12}(x_1-\bar{x})(y_1-\bar{y}) + a_{11}(y_1-\bar{y})^2}{a_{11}a_{22} - a_{12}^2}\right] > 0.$$

This fact appears in later derivations. With Equations (2.1.6) and Equation (2.1.5) the joint p.d.f. of X_1,Y_1 and T is obtained (A.3, Appendix) as

$$f_{X_1,Y_1|T}(\bar{x}_1,\bar{y}_1,a_{11},a_{12},a_{22}) \tag{2.1.7}$$

$$= \frac{(n-1)\Gamma\left(\frac{n-1}{2}\right)}{n\pi\Gamma\left(\frac{n-3}{2}\right)|A|^{\frac{1}{2}}}\left[1 - \frac{n}{n-1}\,\frac{a_{22}(x_1-\bar{x})^2 - 2a_{12}(x_1-\bar{x})(y_1-\bar{y}) + a_{11}(y_1-\bar{y})^2}{|A|}\right]^{\frac{1}{2}(n-5)}$$

$$\text{if } a_{22}(x_1-\bar{x})^2-2a_{12}(x_1-\bar{x})(y_1-\bar{y})+a_{11}(y_1-\bar{y})^2 < \frac{(n-1)|A|}{n},$$

$= 0$ otherwise

where $|A| = a_{11}a_{22}-a_{12}^2$

Then $\hat{g}(\theta) = \hat{F}(h,k) = [E\ \tilde{g}(\theta)|T]$ gives (A.4, Appendix)

$$\hat{F}(h,k) = 1 \qquad h > \beta(k),\ y > \beta$$

$$= \int_{\alpha}^{\beta}\int_{\alpha(k)}^{h} f(\cdot)dxdy \qquad \alpha(k) < h < \beta(k),\ k \geq \beta$$

$$= \int_{\alpha}^{k}\int_{\alpha(k)}^{\beta(k)} f(\cdot)dxdy \qquad h \geq \alpha(k),\ \alpha < k < \beta$$

$$= \int_{\alpha}^{k}\int_{\alpha(k)}^{h} f(\cdot)dxdy \qquad \alpha(h) < h < \beta(k),\ \alpha < k < \beta$$

$$= 0 \qquad \text{otherwise}$$

where f(•) is Equation (2.1.7) and where

$$\alpha(k) = \bar{x} + \frac{a_{12}(k-\bar{y}) - \sqrt{|A|\left[\frac{(n-1)a_{22}}{n} - (k-\bar{y})^2\right]}}{a_{22}},$$

$$\rho(k) = \bar{x} + \frac{a_{12}(k-\bar{y}) + \sqrt{|A|\left[\frac{(n-1)a_{22}}{n} - (k-\bar{y})^2\right]}}{a_{22}}, \tag{2.1.8}$$

$$\alpha = \bar{y} - \sqrt{\frac{(n-1)a_{22}}{n}},$$

$$\beta = \bar{y} + \sqrt{\frac{(n-1)a_{22}}{n}}.$$

In order to make Estimate (2.1.8) more useful it is to be transformed into some special distribution for which tables are available. Let

$$Z_1 = \left(\frac{|A|}{a_{11}}\right)^{\frac{1}{2}} \frac{[n(n-3)]^{\frac{1}{2}}(X_1-\bar{X})}{\left\{(n-1)|A| - n\left[a_{22}(X_1-\bar{X})^2 - 2a_{12}(X_1-\bar{X})(Y_1-\bar{Y}) + a_{11}(Y_1-\bar{Y})^2\right]\right\}^{\frac{1}{2}}},$$

(2.1.9)

$$Z_2 = \left(\frac{|A|}{a_{22}}\right)^{\frac{1}{2}} \frac{[n(n-3)]^{\frac{1}{2}}(Y_1-\bar{Y})}{\left\{(n-1)|A| - n\left[a_{22}(X_1-\bar{X})^2 - 2a_{12}(X_1-\bar{X})(Y_1-\bar{Y}) + a_{11}(Y_1-\bar{Y})^2\right]\right\}^{\frac{1}{2}}},$$

Using this transformation Equation (2.1.7) takes the form (A.5, Appendix)

$$f(Z_1,Z_2) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\pi(n-3)\Gamma\left(\frac{n-3}{2}\right)|R|^{\frac{1}{2}}} \left(1 + \frac{1}{n-3}\, \underset{\sim}{Z}'R^{-1}\underset{\sim}{Z}\right)^{-\frac{1}{2}(n-1)}, \quad \underset{\sim}{Z} \in R^2 \tag{2.1.10}$$

where

$$\underset{\sim}{Z}' = (Z_1, Z_2), \quad R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}, \quad r = \frac{a_{12}}{\sqrt{a_{11}a_{22}}},$$

and Estimate (2.1.8) becomes

$$\hat{F}(\infty,\infty) = 1$$

$$F(h,k) = \int_{-\infty}^{\infty}\int_{-\infty}^{Z_1(h,k)} \frac{1}{2\pi(1-r^2)^{\frac{1}{2}}}\left(1 + \frac{1}{n-3}\underset{\sim}{Z}'R^{-1}\underset{\sim}{Z}\right)^{-\frac{1}{2}(n-1)} dZ_1 dZ_2$$

$$Z_2 \in R, \; -\infty < Z_1 < Z_1(h,k)$$

$$= \int_{-\infty}^{Z_2(h,k)}\int_{-\infty}^{\infty} \frac{1}{2\pi(1-r^2)^{\frac{1}{2}}}\left(1 + \frac{1}{n-3}\underset{\sim}{Z}'R^{-1}\underset{\sim}{Z}\right)^{-\frac{1}{2}(n-1)} dZ_1 dZ_2 \tag{2.1.11}$$

$$Z_1 \in R, \; -\infty < Z_2 < Z_2(h,k)$$

$$= \int_{-\infty}^{Z_2(h,k)}\int_{-\infty}^{Z_1(h,k)} \frac{1}{2\pi(1-r^2)^{\frac{1}{2}}}\left(1 + \frac{1}{n-3}\underset{\sim}{Z}'R^{-1}\underset{\sim}{Z}\right)^{-\frac{1}{2}(n-1)} dZ_1 dZ_2$$

$$-\infty < Z_1 < Z_1(h,k), \; -\infty < Z_2 < Z_2(h,k)$$

$$= 0 \quad \text{otherwise}$$

where

$$Z_1(h,k) = \left(\frac{|A|}{a_{11}}\right)^{\frac{1}{2}} \frac{[n(n-3)]^{\frac{1}{2}}(h-\bar{x})}{\left\{(n-1)\ A - n\left[a_{22}(h-\bar{x})^2 - 2a_{12}(h-\bar{x})(k-\bar{y}) + a_{11}(k-y)^2\right]\right\}^{\frac{1}{2}}}$$

$$(2.1.12)$$

$$Z_2(h,k) = \left(\frac{|A|}{a_{22}}\right)^{\frac{1}{2}} \frac{[n(n-3)]^{\frac{1}{2}}(k-\bar{y})}{\left\{(n-1)\ A - n\left[a_{22}(h-\bar{x})^2 - 2a_{12}(h-\bar{x})(k-\bar{y}) + a_{11}(k-\bar{y})^2\right]\right\}^{\frac{1}{2}}}$$

Equation (2.1.10) is a bivariate t density with $n - 3$ degrees of freedom. The desired probabilities may be found from tables by Dunnett and Sobel [6].

In the above derivations it is assumed that the sample size is at least four. This fact first appears in Equation (2.1.5). Best estimates of $F(h,k)$ for samples of size one and two are investigated in this section. The case $n = 3$ is discussed in the Summary.

For $n = 1$, let (X_1,Y_1) be a random sample of size one from the population with a density given by Equation (2.1.1). $T = \{X_1, Y_1, X_1^2, Y_1^2, X_1 \cdot Y_1\}$ is complete sufficient for μ_x, μ_y, σ_x^2, σ_y^2, and ρ (A.6, Appendix). Let $\tilde{g}(\theta)$ be the same unbiased estimator of $F(h,k)$. Then

$$E[\tilde{g}(\theta)|T] = P(X_1 \le h,\ Y_1 \le k | x_1, y_1)$$

$$= 1 \qquad x_1 \le h,\ y_1 \le k$$

$$= 0 \qquad \text{otherwise.}$$

For $n = 2$, let $\{(X_1,Y_1),\ (X_2,Y_2)\}$ be a random sample of size two from this population. $T = \{X_1 + X_2,\ Y_1 + Y_2,\ X_1^2 + X_2^2,\ Y_1^2 + Y_2^2,\ X_1Y_1 + X_2Y_2\}$ is complete sufficient for μ_x, μ_y, σ_x^2, σ_y^2, and ρ. Let $\tilde{g}(\theta)$ be similarly defined. Then (A.7, Appendix)

$$E[\tilde{g}(\theta)|T] = P(X_1 < h,\ Y_1 < k|T)$$

$$= 1 \quad h \geq \ell_2,\ k \geq m_2$$

$$= \frac{1}{2} \quad \ell_1 < h < \ell_2,\ k \geq m_2$$

$$= \frac{1}{2} \quad \ell_1 \leq h,\ m_1 < k < m_2$$

$$= \frac{1}{2} \quad \ell_1 < h < \ell_2,\ m_1 < k < m_2$$

$$= 1 \quad \text{otherwise}$$

where ℓ_1, ℓ_2, m_1, m_2 are values of X_1, X_2, Y_1, Y_2 to the solution of $T = t$.

Example 1

The shear strength Y and the well diameter X of spot welds have a bivariate normal distribution. Specifications call for X greater than 0.1978 inches and Y greater than 1103.98 pounds. A random sample size of 33 pair observations of a population of such wells yield

$$\bar{x} = 0.1978 \qquad \bar{y} = 1103.98 \qquad r = 0.5003$$

$$s_x = 0.1518 \qquad s_{xy} = 1.8330 \qquad s_y = 24.1350$$

what fraction of this population of welds having X and Y less than the specification?

From Equations (2.11) and (2.12) the estimate of the probability is given by

$$\int_{-\infty}^{0}\int_{-\infty}^{0} \frac{1}{2\pi(1-r^2)^{\frac{1}{2}}} \left(1 + \frac{1}{30} \underset{\sim}{Z}'R^{-1}\underset{\sim}{Z}\right)^{-16} dZ_1 dZ_2$$

which is a bivariate t distribution with 30 degrees of freedom. Its value from Dunnet and Sobel [6] is found to be 0.33333. The use of the cumulative bivariate normal by substituting the estimates for the parameters also gives a value of 0.33333.

2.2 ρ Known, μ_x, μ_y Unknown, σ_x^2, σ_y^2 Known

Let $\{(X_1,Y_1), \ldots, (X_n,Y_n)\}$ be a random sample of size n from a population with density

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y(1-\ ^2)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_y}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\} \tag{2.2.1}$$

where μ_x and μ_y are unknown. $T = \{\overline{X},\overline{Y}\}$ is complete sufficient for μ_x and μ_y. The p.d.f. of T is

$$f_T(\overline{x},\overline{y}) = \frac{n}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{n}{2(1-\rho^2)}\left[\left(\frac{\overline{x}-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{\overline{x}-\mu_x}{\sigma_x}\right)\left(\frac{\overline{y}-\mu_y}{\sigma_y}\right)\right.\right.$$

$$\left.\left. + \left(\frac{\overline{y}-\mu_y}{\sigma_y}\right)^2\right]\right\} \qquad \mu_x, \mu_y \in R, \ \overline{x}, \overline{y} \in R. \tag{2.2.2}$$

Let $\tilde{g}(\theta)$, $\overline{X}^*$, and $\overline{Y}^*$ be similarly defined as in Section 2.1.

$T = \{\overline{X}^*,\overline{Y}^*\}$ is complete sufficient for μ_x and μ_y. The p.d.f. of T^* is

$$f_{T^*}(\overline{x}^*,\overline{y}^*) = \frac{n-1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{n-1}{2(1-\rho^2)}\left[\left(\frac{\overline{x}^*-\mu_x}{\sigma_x}\right)^2 -2\left(\frac{\overline{x}^*-\mu_x}{\sigma_x}\right)\left(\frac{\overline{y}^*-\mu_y}{\sigma_y}\right)\right.\right. \tag{2.2.3}$$

$$\left.\left. + \left(\frac{\overline{y}^*-\mu_y}{\sigma_y}\right)^2\right]\right\} \qquad \mu_x,\ \mu_y \ \varepsilon\ R,\ \overline{x}^*,\ \overline{y}^* \ \varepsilon\ R$$

The joint p.d.f. $f_{X_1,Y_1,T^*}(x_1,y_1,\overline{x}^*,\overline{y}^*)$ of X_1, Y_1, and T^* is just

$$f_{X_1,Y_1,T^*}(x_1,y_1,\overline{x}^*,\overline{y}^*) = f(x_1,y_1)f_{T^*}(\overline{x}^*,\overline{y}^*) \tag{2.2.4}$$

where $f(x_1,y_1)$ is Equation (2.2.1). Let

$$X_1 = X_1, \quad Y_1 = Y_1, \quad \overline{X}^* = \frac{n\overline{X}-X_1}{n-1}, \quad \overline{Y}^* = \frac{n\overline{Y}-Y_1}{n-1} \tag{2.2.5}$$

With Equations (2.2.4) and (2.2.5) the joint p.d.f. $f_{X_1,Y_1,T}(x_1,y_1,\overline{x},\overline{y})$ is obtained; and with Equation (2.2.2) the conditional density of X_1 and Y_1 given T is

$$f_{X_1,Y_1|T}(x_1,y_1|\overline{x},\overline{y})$$

$$= \frac{n}{(n-1)2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{n}{2(n-1)(1-\rho^2)}\left[\left(\frac{x_1-\overline{x}}{\sigma_x}\right)^2 -2\rho\left(\frac{x_1-\overline{x}}{\sigma_x}\right)\left(\frac{y_1-\overline{y}}{\sigma_y}\right)\right.\right.$$

$$\left.\left. + \left(\frac{y_1-\overline{y}}{\sigma_y}\right)^2\right]\right\} \qquad x_1,\ y_1 \ \varepsilon\ R.$$

Then $E[\tilde{g}(\theta)|T]$ gives

$$\hat{F}(h,k) = \int_{-\infty}^{k}\int_{-\infty}^{h} f_{X_1,Y_1|T}(x_1,y_1|\bar{x},\bar{y})dx_1dy_1. \tag{2.2.6}$$

To evaluate this integral let

$$Z_1 = \left(\frac{n}{n-1}\right)^{\frac{1}{2}}\left(\frac{X_1-\bar{X}}{\sigma_x}\right), \quad Z_2 = \left(\frac{n}{n-1}\right)^{\frac{1}{2}}\left(\frac{Y_1-\bar{Y}}{\sigma_y}\right) .$$

Integral (2.2.6) becomes

$$\hat{F}(h,k) = \int_{-\infty}^{Z_2(k)}\int_{-\infty}^{Z_1(h)} \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp\left[-\frac{1}{2(1-\rho^2)}(Z_1^2-2\rho Z_1Z_2+Z_2^2)\right]dZ_1dZ_2 \tag{2.2.7}$$

$$h,\ k \in R$$

where

$$Z_1(h) = \left(\frac{n}{n-1}\right)^{\frac{1}{2}}\left(\frac{h-\bar{x}}{\sigma_x}\right), \quad Z_2(k) = \left(\frac{n}{n-1}\right)^{\frac{1}{2}}\left(\frac{k-\bar{y}}{\sigma_y}\right) .$$

This probability can be readily found from a bivariate normal table.

Estimator (2.2.6) applies for $n \geq 2$. If $n = 1$, let $\{(X_1,Y_1)\}$ be a random sample. $T = \{X_1,Y_1\}$ is complete sufficient for μ_x and μ_y. Then

$$\hat{F}(h,k) = E[\tilde{g}(\theta)|T]$$

$$= P(X_1 \leq h, Y_1 \leq k|x_1,y_1)$$

$$= 1 \quad x_1 \leq h, y_1 \leq k$$

$$= 0 \quad \text{otherwise}.$$

Example 2

In evaluating the drop-out rate in a large university, the office of student testing and evaluation has noted that X, the combined score on the college board examination, and Y, the grade point at the end of the freshman year, have a bivariate normal distribution with $\rho = 0.60$, $\sigma_x = 80.00$, and $\sigma_y = 0.50$. The office has also noted that an X score less than 1220 and a Y score less than 1.8 would result in a drop-out. A grade point of 4.00 correspondes to A. A random sample of size 6 from a freshman class yield an $\overline{x}$ of 1300 and a $\overline{y}$ of 2.30. What would be the drop-out rate of the freshman class?

By Equation (2.2.7) this rate is estimated as

$$\int_{-\infty}^{-1.1}\int_{-\infty}^{-1.1} n_2(0,0,1,1,0.6)dxdy = 0.058,$$

or the drop-out rate is 5.8 percent. The use of the cumulative bivariate normal by substituting $\overline{X}$ and $\overline{Y}$ for μ_x and μ_y gives a rate of

$$\int_{-\infty}^{-1.0}\int_{-\infty}^{-1.0} n_2(0,0,1,1,0.6)dxdy = 0.073,$$

or 7.3 percent.

2.3 $\rho = 0$, μ_x, μ_y Unknown, σ_x^2, σ_y^2 Known

This case parallels the case given in Section 2.2. The best estimate $\hat{F}(h,k)$ of $F(h,k)$ where

$$F(h,k) = \int_{-\infty}^{k} \int_{-\infty}^{h} \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\} dxdy,$$

$$\mu_x,\ \mu_y \in R,\ \sigma_x^2,\ \sigma_y^2 > 0,\ x,y \in R$$

is given by

$$\hat{F}(h,k) = F\left[\left(\frac{n}{n-1}\right)^{\frac{1}{2}} \cdot \frac{h-\bar{x}}{\sigma_x}\right] F\left[\left(\frac{n}{n-1}\right) \cdot \frac{k-\bar{y}}{\sigma_y}\right]. \tag{2.3.1}$$

where F is the cumulative distribution of a standard normal variate.

For n = 1, the best estimate $\hat{F}(h,k)$ of F(h,k) is again either 0 or 1 as given in Section 2.2.

<u>Example 3</u>

If, in Example 2, the correlation between X and Y is zero, what would be the drop-out rate?

By Equation (2.3.1), the rate is approximated by

$$\hat{F}(1220,\ 1.8) = F(-1.1)F(-1.1)$$

$$= .018$$

or 1.8 percent while by using $\bar{X}$ and $\bar{Y}$ for μ_x and μ_y respectively the rate is about 2.5 percent as given by F(-1)F(-1).

2.4 $\rho = 0$, μ_x, μ_y Known, σ_x^2, σ_y^2 Unknown

Let $\{(X_1,Y_1),\ (X_2,Y_2),\ \ldots,\ (X_n,Y_n)\}$ be a random sample of size n from a population with density

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}, \tag{2.4.1}$$

$$\sigma_x^2,\ \sigma_y^2,\ > 0,\ x,y \in R.$$

Let

$$A_{11} = \sum_{i=1}^{n} (X_i-\mu_x)^2, \qquad A_{22} = \sum_{i=1}^{n} (Y_i-\mu_y)^2.$$

$T = \{A_{11}, A_{22}\}$ is complete sufficient for σ_x^2 and σ_y^2 (A.8, Appendix).

Let $U = \frac{A_{11}}{\sigma_x^2}$. U has a Chi-Square distribution with n degrees of freedom.

That is,

$$f(u) = \frac{1}{2^n\Gamma\left(\frac{n}{2}\right)} u^{\frac{n-2}{2}} e^{-\frac{u}{2}}.$$

Noting that $A_{11} = \sigma_x^2 U$, the density of A_{11} is then

$$f(a_{11}) = \frac{1}{2^n\Gamma\left(\frac{n}{2}\right)\sigma_x^2} \left(\frac{a_{11}}{\sigma_x^2}\right)^{\frac{n-2}{2}} e^{-\frac{1}{2}\left(\frac{a_{11}}{\sigma_x^2}\right)},\quad a_{11} \geq 0.$$

Since A_{11} and A_{22} are independent, the density of T is

$$f_T(a_{11}, a_{22}) = \frac{1}{\left[2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)\right]^2 \sigma_x^2\sigma_y^2} \left(\frac{a_{11}}{\sigma_x^2}\right)\cdot\left(\frac{a_{22}}{\sigma_y^2}\right)^{\frac{n-2}{2}} e^{-\frac{1}{2}\left(\frac{a_{11}}{\sigma_x^2} + \frac{a_{22}}{\sigma_y^2}\right)},$$

$$a_{11},\ a_{22} > 0.$$

Let $\tilde{g}(\theta)$ be similarly defined. Let

$$A_{11}^{*} = \sum_{i=2}^{n} (X_i-\mu_x)^2, \quad A_{22}^{*} = \sum_{i=2}^{n} (Y_i-\mu_y)^2$$

$T^{*} = \{A_{11}^{*}, A_{22}^{*}\}$ is complete sufficient for σ_x^2 and σ_y^2. The density of T^{*} is

$$f(a_{11}^{*}, a_{22}^{*}) = \frac{1}{\left[2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)\right]^2 \sigma_x^2 \sigma_y^2} \left(\frac{a_{11}^{*}}{\sigma_x^2} \cdot \frac{a_{22}^{*}}{\sigma_y^2}\right)^{\frac{n-3}{2}} e^{-\frac{1}{2}\left(\frac{a_{11}^{*}}{\sigma_x^2} + \frac{a_{22}^{*}}{\sigma_y^2}\right)}, \tag{2.4.3}$$

$$a_{11},\ a_{22} \geq 0.$$

The joint p.d.f. of X_1, Y_1, and T^{*} is

$$f_{X_1,Y_1,T^{*}}(x_1,y_1,a_{11},a_{22}) = f(x_1,y_1) f_{T^{*}}(a_{11}^{*}, a_{22}^{*}). \tag{2.4.4}$$

Let

$$X_1 = X_1, \quad Y_1 = Y_1,$$

$$A_{11}^{*} = A_{11} - (X_1-\mu_X)^2, \quad A_{22}^{*} = A_{22} - (Y_1-\mu_y)^2 \tag{2.4.5}$$

with Equations (2.2.4) and (2.4.5) $f_{X_1,Y_1,T}(x_1,y_1,a_{11},a_{22})$ is straightforwardly obtained. Then the conditional density of X_1 and Y_1 given T is

$$f_{X_1,Y_1|T}(x_1,y_1|a_{11},a_{22}) \tag{2.4.6}$$

$$= \frac{1}{\beta\left(\frac{1}{2}, \frac{n-1}{2}\right)\sqrt{a_{11}}}\left[1 - \frac{(x_1-\mu_x)^2}{a_{11}}\right]^{\frac{n-3}{2}} \frac{1}{\beta\left(\frac{1}{2}, \frac{n-1}{2}\right)\sqrt{a_{22}}}\left[1 - \frac{(y_1-\mu_y)^2}{a_{22}}\right]^{\frac{n-3}{2}}$$

$$-\sqrt{a_{11}} < x_1-\mu_x < \sqrt{a_{11}}, \quad -\sqrt{a_{22}} < y_1-\mu_y < \sqrt{a_{22}}$$

where β is the beta function. Let

$$W_1 = \frac{(n-1)^{\frac{1}{2}}(X_1-\mu_x)}{\sqrt{a_{11}}} \Bigg/ \left[1 - \frac{(X_1-\mu_x)^2}{a_{11}}\right]^{\frac{1}{2}},$$

$$W_2 = \frac{(n-1)^{\frac{1}{2}}(Y_1-\mu_y)}{\sqrt{a_{22}}} \Bigg/ \left[1 - \frac{(Y_1-\mu_y)^2}{a_{22}}\right]^{\frac{1}{2}}. \tag{2.4.7}$$

The Jacobian of the transformation is (A.8, Appendix)

$$\left|\frac{\partial(x_1,y_1)}{\partial(w_1,w_2)}\right| = \left(\frac{a_{11}}{n-1}\right)^{\frac{1}{2}}\left(1 + \frac{w_1^2}{n-1}\right)^{-\frac{3}{2}}\left(\frac{a_{22}}{n-1}\right)^{\frac{1}{2}}\left(1 + \frac{w_2^2}{n-1}\right)^{-\frac{3}{2}}.$$

Equation (2.4.6) in W_1 and W_2 takes the form

$$f\left(x_1(w_1), y_1(w_2), a_{11}, a_{22}\right) \tag{2.4.8}$$

$$= \frac{1}{(n-1)^{\frac{1}{2}}\beta\left(\frac{1}{2},\frac{n-1}{2}\right)}\left(1 + \frac{w_1^2}{n-1}\right)^{-\frac{n}{2}} \frac{1}{(n-1)^{\frac{1}{2}}\beta\left(\frac{1}{2},\frac{n-1}{2}\right)}\left(1 + \frac{w_2^2}{n-1}\right)^{-\frac{n}{2}},$$

$$w_1, w_2 \in R.$$

Then the best estimate $\hat{F}(h,k)$ of $F(h,k)$ as given by $E\,\tilde{g}(\Theta)|T$ is

$$\hat{F}(h,k) = 1 \quad h > \mu_x + \sqrt{a_{11}}, \quad k > \mu_y + \sqrt{a_{22}} \tag{2.4.9}$$

$$= F_{t,n-1}(w_{10}) \quad \mu_x - \sqrt{a_{11}} < h < \mu_x + \sqrt{a_{11}}, \quad k \geq \mu_y + \sqrt{a_{22}}$$

$$= F_{t,n-1}(w_{20}) \quad h \geq \mu_x + \sqrt{a_{11}}, \quad \mu_y - \sqrt{a_{22}} < k < \mu_y + \sqrt{a_{22}}$$

$$= F_{t,n-1}(w_{10})F_{t,n-1}(w_{20}) \quad \mu_x - \sqrt{a_{11}} < h < \mu_x + \sqrt{a_{11}},$$

$$\mu_y - \sqrt{a_{22}} < k < \mu_y + \sqrt{a_{22}}$$

$$= 0 \quad \text{otherwise}$$

where $F_{t,n-1}$ denotes the cumulative Student's distribution and where

$$w_{10} = (n-1)^{\frac{1}{2}}(h-\mu_x) \Big/ \left[a_{22}-(h-\mu_x)^2\right]^{\frac{1}{2}},$$

$$w_{20} = (n-1)^{\frac{1}{2}}(k-\mu_y) \Big/ \left[a_{11}-(k-\mu_y)^2\right]^{\frac{1}{2}}. \tag{2.4.10}$$

Estimator (2.4.9) applies for $n \geq 2$. For $n = 1$, let $\{(X_1,Y_1)\}$ be a random sample of size one. $T = \{(X_1-\mu_x)^2, (Y_1-\mu_y)^2\}$ is complete sufficient for σ_x^2 and σ_y^2. Let $\tilde{g}(\theta)$ be similarly defined as before. Then (A.8, Appendix)

$$E[\tilde{g}(\theta)|T] = P\left(X_1<h, \quad Y_1<k \mid (x_1-\mu_x)^2, (y_1-\mu_y)^2\right)$$

$$= 1 \quad h > \ell_2, \; k > m_2$$

$$= \frac{1}{2} \quad h \geq \ell_2,\ m_1 < k < m_2$$

$$= \frac{1}{2} \quad \ell_1 < h < \ell_2,\ k \geq m_2$$

$$= \frac{1}{4} \quad \ell_1 < h < \ell_2,\ m_1 < k < m_2$$

$$= 0 \quad \text{otherwise}$$

where (ℓ_1,m_1), (ℓ_1,m_2), (ℓ_2,m_1) (ℓ_2,m_2) are values of (X_1,Y_1) to the solution of $(X_1-\mu_x)^2 = t_1$, and $(Y_1=\mu_y)^2 = t_2$. For convenience it is assumed that $\ell_1 < \ell_2$ and $m_1 < m_2$.

Example 4

Suppose that in Example 2 $\mu_x = 1300$, $\mu_y = 2.30$, $\rho = 0$, σ_x and σ_y are unknown. A random sample of size ten gives $s_x^2 = 6400$ and $s_y^2 = 0.25$. What would be the drop-out rate?

By Equations (2.4.10) w_{10} and w_{20} are computed as -1.06 and -1.06. The estimate of the drop-out rate is $\left[F_{t,32}(-1.06)\right]^2$ as given by Equation (2.4.9). By linear interpolation from the Table [3], $\left[F_{t,9}(-1.06)\right]^2$ is about 0.0293. The drop-out rate is about 2.5 percent if the cumulative bivariate normal distribution is used.

2.5 $\rho = 0$, μ_y, σ_x^2 Known, μ_x, σ_y^2 Unknown

Let (X_1,Y_1), (X_2,Y_2), ..., (X_n,Y_n) be a random sample of size n from a population with density

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}. \tag{2.5.1}$$

$T = \{\overline{X}, A_{22}\}$ is complete sufficient for μ_x and σ_y^2. The density of T is

$$f_T(\overline{x}, a_{22}) = \frac{\sqrt{n}}{2^{\frac{n}{2}}\sqrt{2\pi}\,\Gamma\left(\frac{n}{2}\right)\sigma_x\sigma_y^2}\left(\frac{a_{22}}{\sigma_y^2}\right)^{\frac{n-2}{2}} \exp\left\{-\frac{1}{2}\left[\frac{a_{22}}{\sigma_y^2} + n\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right]\right\}, \tag{2.5.2}$$

$$\overline{x} \in R,\ a_{22} \geq 0.$$

Also, $T^* = \{\overline{X}^*, A_{22}^*\}$ is complete sufficient for μ_x and σ_y^2. The density $f_{T^*}(\overline{x}^*, a_{22}^*)$ of T^* patterns after Equation (2.5.2) with n being n - 1. The joint p.d.f. of X_1, Y_1 and T^* is

$$f_{X_1,Y_1,T^*}(x_1,y_1,\overline{x}^*,a_{11}^*) = f(x_1,y_1)f_{T^*}(\overline{x}^*,a_{22}^*). \tag{2.5.3}$$

Let

$$X_1 = X_1,\quad Y_1 = Y_1,\quad \overline{X}^* = \frac{n\overline{X}-X_1}{n-1},\quad A_{22}^* = A_{22}-(Y_1-\mu_y)^2. \tag{2.5.4}$$

With Equations (2.5.4) and (2.5.3) the joint p.d.f. $f_{X_1,Y_1,T}(x_1,y_1,\overline{x},a_{22})$ is easily obtained. Then

$$f_{X_1,Y_1|T}(x_1,y_1|\bar{x},a_{22}) = f_{X_1,Y_1,T}(x_1,y_1,\bar{x},a_{22}) \Big/ f_T(\bar{x},a_{22}) \qquad (2.5.5)$$

$$= \frac{n}{\sqrt{(n-1)2\pi}\sigma_x} \exp\left[-\frac{n}{2(n-1)}\left(\frac{x_1-\bar{x}}{\sigma_x}\right)^2\right] \frac{1}{\beta\left(\frac{1}{2},\frac{n-1}{2}\right)\sqrt{a_{22}}}\left[1-\frac{(y_1-\mu_y)^2}{a_{22}}\right]^{\frac{n-3}{2}},$$

$$x_1 \in R,\ -\sqrt{a_{22}} < y_1-\mu_y < \sqrt{a_{22}}.$$

Let

$$Z = \left(\frac{n}{n-1}\right)^{\frac{1}{2}}\left(\frac{X_1-\bar{X}}{\sigma_x}\right), \quad W = \frac{(n-1)^{\frac{1}{2}}(Y_1-\mu_y)}{\sqrt{a_{22}}} \Bigg/ \left[1-\frac{(y_1-\mu_y)^2}{a_{22}}\right]^{\frac{1}{2}}. \qquad (2.5.6)$$

The Jacobian of the transformation is

$$\left|\frac{\partial(x_1,y_1)}{\partial(Z,w)}\right| = \frac{\sigma_x\sqrt{a_{22}}}{n}\left(1+\frac{w^2}{n-1}\right)^{-\frac{3}{2}}.$$

Equation (2.5.5) is transformed to

$$f(Z,w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} \frac{1}{(n-1)^{\frac{1}{2}}\beta\left(\frac{1}{2},\frac{n-1}{2}\right)}\left(1+\frac{w^2}{n-1}\right)^{\frac{n}{2}}, \ Z \in R,\ w \in R.$$

Then $E\left[\tilde{g}(\theta)|T\right]$ gives

$$F(h,k) = 0 \qquad k < \mu_y - \sqrt{a_{22}}$$

$$= F\left[\left(\frac{n}{n-1}\right)^{\frac{1}{2}} \frac{h-x}{\sigma_x}\right] \qquad k > \mu_y + \sqrt{a_{22}} \tag{2.5.7}$$

$$= F\left[\left(\frac{n}{n-1}\right)^{\frac{1}{2}} \frac{h-\bar{x}}{\sigma_x}\right] F_{t,n-1}(w_0) \qquad \text{if } \mu_y - \sqrt{a_{22}} < k < \mu_y + \sqrt{a_{22}}$$

where F is the standard normal distribution and $F_{t,n-1}$ is the Student's distribution with n - 1 degrees of freedom where

$$w_0 = (n-1)^{\frac{1}{2}}(k-\mu_y)/\left[a_{22}-(y_1-\mu_y)^2\right]^{\frac{1}{2}}.$$

Equation (2.5.7) is valid for $n \geq 2$ for $n = 1$, $T = \{X_1, (Y_1-\mu_y)^2\}$ is complete sufficient for μ_x and σ_y^2. Let (ℓ_1, m_1) and (ℓ_2, m_2) be the values of (X_1, Y_1) to the solution of $T = t$, and $m_1 < m_2$. Then $E\, \tilde{g}(\theta)|T$ gives

$$\hat{F}(h,k) = 1 \qquad h > \ell_1,\ k > m_2$$

$$= \frac{1}{2} \qquad h > \ell_1,\ m_1 < k < m_2$$

$$= 0 \qquad \text{otherwise}$$

<u>Example 5</u>

If in Example 2 $\mu_y = 2.3$, $\sigma_x^2 = 6400$, $\rho = 0$, μ_x and σ_y^2 unknown, a random sample of size ten gives an $\overline{X}$ of 1300 and s_y^2 of 0.25. What would be the drop-out rate?

Equations (2.5.6) and (2.5.7) give $F(-1.05)F_{t,9}(-1.06)$ $= (.1469)(.1713) = 0.025$ as the estimate. The use of a cumulative bivariate normal also gives a drop-out rate of 2.5 percent.

2.6 $\rho = 0$, μ_x, μ_y, σ_x^2, σ_y^2 Unknown

Let $\{(X_1,Y_1), \ldots, (X_n,Y_n)\}$ be a random sample from a population with density

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\} \tag{2.6.1}$$

$$\mu_x,\ \mu_y \in R,\ \sigma_x,\ \sigma_y > 0;\ x,\ y \in R.$$

Let

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

$$A_{11} = \sum_{i=1}^{n}(X_i-\overline{X})^2, \qquad A_{22} = \sum_{i=1}^{n}(Y_i-\overline{Y})^2.$$

$T = \{\overline{X}, \overline{Y}, A_{11}, A_{22}\}$ is complete sufficient for μ_x, μ_y, σ_x^2, σ_y^2. The density of T is

$$f_T(\overline{x},\overline{y},a_{11},a_{22}) = \frac{n}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{n}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}\cdot$$

$$\frac{1}{2^{n-1}\left[\Gamma\left(\frac{n-1}{2}\right)\right]^2\sigma_x^2\sigma_y^2}\left[\left(\frac{a_{11}}{\sigma_x^2}\right)\cdot\left(\frac{a_{22}}{\sigma_y^2}\right)\right]^{\frac{n-3}{2}} \exp\left[-\frac{1}{2}\left(\frac{a_{11}}{\sigma_x^2} + \frac{a_{22}}{\sigma_y^2}\right)\right],$$

$$x,\ y \in R;\ a_{11},\ a_{22} \geq 0.$$

Let

$$\overline{X}^* = \frac{1}{n-1}\sum_{i=2}^{n} X_i, \quad \overline{Y}^* = \frac{1}{n-1}\sum_{i=2}^{n} Y_i,$$

$$A_{11}^* = \sum_{i=2}^{n}(X_i-\overline{X}^*)^2, \quad A_{22}^* = \sum_{i=2}^{n}(Y_i-\overline{Y}^*)^2.$$

$T^* = \{\overline{X}^*, \overline{Y}^*, A_{11}^*, A_{22}^*\}$ is complete sufficient for μ_x, μ_y, σ_x^2, and σ_y^2. The density $f_{T^*}(\overline{x}^*,\overline{y}^*,a_{11}^*,a_{22}^*)$ is obtained from Equation (2.6.2) with n being replaced by n - 1. Then the joint p.d.f. of X_1, Y_1 and T^* is

$$f_{X_1,Y_1,T^*}(x_1,y_1,\overline{x}^*,\overline{y}^*,a_{11}^*,a_{22}^*) = f(x_1,y_1)f_{T^*}(\overline{x}^*,\overline{y}^*,a_{11}^*,a_{22}^*). \qquad (2.6.3)$$

Let

$$X_1 = X_1, \quad Y_1 = Y_1, \quad \overline{X}^* = \frac{n\overline{X}-X_1}{n-1}, \quad \overline{Y}^* = \frac{n\overline{Y}-Y_1}{n-1},$$

$$(2.6.4)$$

$$A_{11}^* = A_{11} - \frac{n}{n-1}(X_1-\overline{X})^2, \quad A_{22}^* = A_{22} - \frac{n}{n-1}(Y_1-\overline{Y})^2.$$

With Equations (2.6.4) and (2.6.3) the joint p.d.f. $f_{X_1,Y_1,T}(x_1,y_1,\overline{x},\overline{y},a_{11},a_{22})$ is obtained. Then the conditional density of X_1 and Y_1 given T is

$$f_{X_1,Y_1|T}(x_1,y_1|\overline{x},\overline{y},a_{11},a_{22}) \tag{2.6.5}$$

$$= \frac{n}{n-1}\left[\frac{\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}\right]^2 \frac{1}{\sqrt{a_{11}a_{22}}}\left[1-\frac{(x_1-\overline{x})^2}{a_{11}}\right]^{\frac{n-4}{2}}\left[1-\frac{(y_1-\overline{y})^2}{a_{22}}\right]^{\frac{n-4}{2}},$$

$$x_1,\ y_1 \ \varepsilon\ R;\ -\sqrt{a_{11}} < \frac{n}{n-1}(x_1-\overline{x}) < \sqrt{a_{11}},\ -\sqrt{a_{22}} < \frac{n}{n-1}(y_1-\overline{y}) < \sqrt{a_{22}}.$$

Let

$$W_1 = \frac{\left[n(n-2)\right]^{\frac{1}{2}}(X_1-\overline{X})}{\left[(n-1)a_{11}-n(X_1-\overline{X})^2\right]^{\frac{1}{2}}},\qquad W_2 = \frac{\left[n(n-2)\right]^{\frac{1}{2}}(Y_1-\overline{Y})}{\left[(n-1)a_{22}-n(Y_1-\overline{Y})^2\right]^{\frac{1}{2}}}.$$

The Jacobian is

$$\left|\frac{\partial(x_1,y_1)}{\partial(w_1,w_2)}\right| = \frac{(n-1)\sqrt{a_{11}a_{22}}}{n(n-2)}\left(1+\frac{w_1^2}{n-2}\right)^{-\frac{3}{2}}\left(1+\frac{w_2^2}{n-2}\right)^{-\frac{3}{2}}.$$

Equation (2.6.5) in w_1 and w_2 is

$$f(w_1,w_2) = \frac{1}{(n-2)^{\frac{1}{2}}\beta\left(\frac{1}{2},\frac{n-2}{2}\right)}\left(1+\frac{w_2^2}{n-2}\right)^{-\frac{n-1}{2}}\frac{1}{(n-2)^{\frac{1}{2}}\beta\left(\frac{1}{2},\frac{n-2}{2}\right)}\left(1+\frac{w_2^2}{n-2}\right)^{-\frac{n-1}{2}}$$

$$w_1,\ w_2\ \varepsilon\ R,$$

where W_1 and W_2 are the Student's variates with $n - 2$ degrees of freedom.

Hence $E[\tilde{g}(\theta)|T]$ gives

$$
\begin{aligned}
F(h,k) &= 1 \quad h > \overline{x} + \sqrt{(n-1)a_{11}/n}, \quad k > \overline{y} + \sqrt{(n-1)a_{22}/n} \\
&= F_{t,n-2}(w_{10}) \quad \overline{x} - \sqrt{(n-1)a_{11}/n} < h < \overline{x} + \sqrt{(n-1)a_{22}/n}\ , \\
&\qquad k > \overline{y} + \sqrt{(n-1)a_{22}/n} \\
&= F_{t,n-2}(w_{20}) \quad h > \overline{x} + \sqrt{(n-1)a_{11}/n}\ , \\
&\qquad y < \sqrt{(n-1)a_{22}/n} < k < \overline{y} + \sqrt{(n-1)a_{22}/n} \\
&= F_{t,n-2}(w_{10})F_{t,n-2}(w_{20}) \quad \overline{x} - \sqrt{(n-1)a_{11}/n} < h < x + \sqrt{(n-1)a_{11}/n}, \\
&\qquad \overline{y} - \sqrt{(n-1)a_{22}/n} < k < \overline{y} + \sqrt{(n-1)a_{22}/n} \\
&= 0 \quad \text{otherwise}
\end{aligned}
\tag{2.6.6}
$$

where

$$
\begin{aligned}
w_{10} &= \left[n(n-2)\right]^{\frac{1}{2}}(h-\overline{x}) \Big/ \left[(n-1)a_{11}-n(h-\overline{x})^2\right]^{\frac{1}{2}} \\
w_{20} &= \left[n(n-2)\right]^{\frac{1}{2}}(k-\overline{y}) \Big/ \left[(n-1)a_{22}-n(k-\overline{y})^2\right]^{\frac{1}{2}}
\end{aligned}
\tag{2.6.7}
$$

Estimate (2.6.6) applies for $n \geq 3$. For $n = 1$, $T = \{X_1, Y_1, X_1^2, Y_1^2\}$ is complete sufficient for μ_x, μ_y, σ_x^2, σ_y^2. Then

$$
\begin{aligned}
\hat{F}(h,k) &= E\left[\tilde{g}(\theta)|T\right] \\
&= E\left[\tilde{g}(\theta)|x_1, y_1, x_1^2, y_1^2\right]
\end{aligned}
$$

$$= E\left[\tilde{g}(\theta)\,|\,x_1,y_1\right]$$

$$= P\left[X_1 < h,\ Y_1 < k\,|\,x_1,y_1\right]$$

$$= 1 \quad \text{if } x_1 \le h \text{ and } y_1 \le k,$$

$$= 0 \quad \text{otherwise.}$$

For $n = 2$, $T = \{X_1+X_2,\ Y_1+Y_2,\ X_1^2,+X_2^2,\ Y_1^2+Y_2^2\}$ is complete sufficient for μ_x, μ_y, σ_x^2 and σ_y^2. Given $T = t$, the solution set in terms of (X_1,Y_1,X_2,Y_2) consists of four points. Let them be denoted by (ℓ_1,m_1,ℓ_2,m_2), (ℓ_2,m_2,ℓ_1,m_1), and (ℓ_1,m_2,ℓ_2,m_1). For convenience let $\ell_1 < \ell_2$ and $m_1 < m_2$. Then

$$\hat{F}(h,k) = E\left[\tilde{g}(\theta)\,|\,T\right]$$

$$= P\left(X_1<h, Y_1<k\,|\,x_1+x_2, y_1+y_2, x_1^2+x_2^2, y_1^2+y_2^2\right)$$

$$= P\left[X_1<h, Y_1<k\,|\,(\ell_1,m_1,\ell_2,m_2),(\ell_2,m_2,\ell_1,m_1),(\ell_2,m_1,\ell_1,m_2),\right.$$
$$\left.(\ell_1,m_2,\ell_2,m_1)\right]$$

$$= 1 \quad h \ge \ell_2,\ k \ge m_2$$

$$= \frac{1}{2} \quad \ell_1 < h < \ell_2,\ k \ge m_2$$

$$= \frac{1}{2} \quad h \ge \ell_2,\ m_1 < k < m_2$$

$$= \frac{1}{4} \quad \ell_1 < h < \ell_2,\ m_1 < k < m_2$$

$= 0$ otherwise.

Example 6

Suppose that in Example 2 $\rho = 0$ and all other parameters are unknown. If a random sample of size 10 gives $\overline{x} = 1300$, $\overline{y} = 2.3$, $s_x^2 = 6400$, and $s_y^2 = 0.25$, what would be the drop-out rate?

Equations (2.6.7) give $w_{10} = -1.06$ and $w_{20} = -1.06$. Equation (2.6.6) gives $\left[F_{t,8}(-1.06)\right]^2$ as the estimate of this rate. By linear interpolation from the Table [3], $\left[F_{t,8}(-1.06)\right]^2$ is about .029. The drop-out rate is about 2.9 percent as compared to 2.5 percent if the cumulative bivariate normal is used.

CHAPTER III

EXTENSION TO A p-VARIATE NORMAL DISTRIBUTION

Extension is developed in this chapter to obtain the best estimate of $F(h_1, \ldots, h_p)$, where F is the cumulative distribution of a p-variate normal variable.

3.1 $\underset{\sim}{\mu}$, Σ Unknown

Let

$$\underset{\sim}{X} = \begin{pmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_p \end{pmatrix}$$

be a p x 1 random vector with density

$$f(\underset{\sim}{x}) = f(x_1, \ldots, x_p) = \frac{1}{(2\mu)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})\right], \qquad (3.1.1)$$

$\underset{\sim}{\mu} = (\mu_1, \ldots, \mu_p)' \varepsilon R^p$, Σ positive definite.

Let $(\underset{\sim}{X}^{(1)}, \underset{\sim}{X}^{(2)}, \ldots, \underset{\sim}{X}^{(n)})$ be a random sample of size n from a population with density Equation (3.1.1). Let

$$\underset{\sim}{\overline{X}} = \frac{1}{n} \sum_{k=1}^{n} X^{(k)} = \begin{pmatrix} \frac{1}{n} \sum_{k=1}^{n} X_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ \frac{1}{n} \sum_{k=1}^{n} X_p^{(k)} \end{pmatrix} = \begin{pmatrix} \overline{X}_1 \\ \cdot \\ \cdot \\ \cdot \\ \overline{X}_p \end{pmatrix}$$

and

$$A = \sum_{k=1}^{n} (\underset{\sim}{X}^{(k)} - \underset{\sim}{\overline{X}})(\underset{\sim}{X}^{(k)} - \underset{\sim}{\overline{X}})'.$$

The densities of $\underset{\sim}{\overline{X}}$ and A are respectively

$$f_{\underset{\sim}{\overline{X}}}(\underset{\sim}{\overline{x}}) = \frac{n^{\frac{p}{2}}}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{n}{2}(\underset{\sim}{\overline{x}}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{\overline{x}}-\underset{\sim}{\mu})\right], \ \underset{\sim}{x}, \ \underset{\sim}{\mu} \ \varepsilon \ R^p \tag{3.1.2}$$

$$f_A(A) = \frac{|A|^{\frac{1}{2}(n-2-p)} \exp\left(-\frac{1}{2} \operatorname{tr}\Sigma^{-1}A\right)}{2^{\frac{1}{2}(n-1)p} \pi^{\frac{1}{4}(p-1)p} |\Sigma|^{\frac{1}{2}(n-1)} \prod_{i=1}^{p} \Gamma\left[\frac{1}{2}(n-i)\right]} . \tag{3.1.3}$$

$T = \{\underset{\sim}{\overline{X}}, A\}$ is complete sufficient for $\underset{\sim}{\mu}$ and Σ (A.9, Appendix). Since $\underset{\sim}{\overline{X}}$ and A are stochastically independent the density of T is the product of Equations (3.1.2) and (3.1.3). That is,

$$f_T(\underset{\sim}{\overline{x}}, a) = f_{\underset{\sim}{\overline{X}}}(\underset{\sim}{\overline{x}}) f_A(a). \tag{3.1.4}$$

Partition the sample $\{\underset{\sim}{X}^{(1)}, \underset{\sim}{X}^{(2)}, \ldots, \underset{\sim}{X}^{(n)}\}$ into two subsamples $\{\underset{\sim}{X}^{(1)}\}$ and $\{\underset{\sim}{X}^{(2)}, \ldots, \underset{\sim}{X}^{(n)}\}$. Let

$$\tilde{g}(\theta) = 1 \quad \text{if } X_1^{(1)} < h_1,\ X_2^{(1)} < h_2,\ \ldots,\ \text{and } X_p^{(1)} < h_p,$$

$$= 0 \quad \text{otherwise}$$

where θ denotes the parameter space $\underset{\sim}{\mu}$ and $\underset{\sim}{\Sigma}$. $\tilde{g}(\theta)$ is unbiased for $F(h_1, \ldots, h_p)$. Let

$$\underset{\sim}{\overline{X}}^* = \frac{1}{n-1} \sum_{k=2}^{n} \underset{\sim}{X}^{(k)} = \begin{pmatrix} \frac{1}{n-1} \sum_{k=2}^{n} X_1^{(k)} \\ \cdot \\ \cdot \\ \cdot \\ \frac{1}{n-1} \sum_{k=2}^{n} X_p^{(k)} \end{pmatrix},$$

and

$$A^* = \sum_{k=2}^{n} (\underset{\sim}{X}^{(k)} - \underset{\sim}{\overline{X}}^*)(\underset{\sim}{X}^{(k)} - \underset{\sim}{\overline{X}}^*)'.$$

The densities $f_{\underset{\sim}{\overline{X}}^*}(\overline{x}^*)$ and $f_{A^*}(a^*)$ of $\overline{X}^*$ and A^* are respectively obtained from Equations (3.1.2) and (3.1.3) with n being replaced by n - 1. The joint density of $\underset{\sim}{X}^{(1)}$ and T^* is

$$f_{\underset{\sim}{X}^{(1)},T^*}(\underset{\sim}{x}^{(1)}, \underset{\sim}{\overline{x}}^*, a^*) = f(\underset{\sim}{x}^{(1)}) f_{\underset{\sim}{\overline{X}}^*}(\underset{\sim}{\overline{x}}^*) f_{A^*}(a^*). \tag{3.1.5}$$

Let

$$\underset{\sim}{X}^{(1)} = \underset{\sim}{X}^{(1)}, \quad \overline{X}^* = \frac{1}{n-1}(n\overline{\underset{\sim}{X}} - \underset{\sim}{X}^{(1)})$$
$$A^* = A - \frac{1}{n-1}(\underset{\sim}{X} - \underset{\sim}{X}^{(1)})(\underset{\sim}{X} - \underset{\sim}{X}^{(1)})'. \tag{3.1.6}$$

The Jacobian of the transformation is (A.10, Appendix)

$$\left|\frac{\partial(\underset{\sim}{x}^{(1)},\overline{x}^*,a^*)}{\partial(\underset{\sim}{x}^{(1)},\overline{\underset{\sim}{x}},a)}\right| = \left(\frac{n}{n-1}\right)^p. \tag{3.1.7}$$

Equations (3.1.5), (3.1.6), and (3.1.7) give the joint density $f_{\underset{\sim}{X}^{(1)},T}(\underset{\sim}{x}^{(1)},\overline{\underset{\sim}{x}},a)$ of $\underset{\sim}{X}^{(1)}$ and T. The conditional density of $\underset{\sim}{X}^{(1)}$ given $\underset{\sim}{X}$ and A is (A.11, Appendix)

$$f_{\underset{\sim}{X}^{(1)}|\overline{\underset{\sim}{X}},A}(\underset{\sim}{x}^{(1)}\ \overline{\underset{\sim}{x}},a) \tag{3.1.8}$$

$$= \left(\frac{n}{n-1}\right)^{\frac{p}{2}} \frac{\Gamma\left(\frac{n}{2}\right)}{\pi^{\frac{p}{2}}\Gamma\left(\frac{n-p}{2}\right)|A|^{\frac{1}{2}}}\left[1 - \frac{n}{n-1}(\underset{\sim}{x}^{(1)} - \overline{\underset{\sim}{x}})'A^{-1}(\underset{\sim}{x}^{(1)} - \overline{\underset{\sim}{x}})\right]^{\frac{1}{2}(n-3-p)},$$

$$\text{when } (\underset{\sim}{x}^{(1)} - \overline{\underset{\sim}{x}})'A^{-1}(\underset{\sim}{x}^{(1)} - \overline{\underset{\sim}{x}}) < \frac{n-1}{n}$$

$= 0$ otherwise.

Then $E[\tilde{g}(\theta)|T]$ gives

$$F(h_1, \ldots, h_p) = \int_{-\infty}^{h_p} \cdots \int_{-\infty}^{h_1} f_{\underset{\sim}{X}^{(1)}\ T}(\underset{\sim}{x}^{(1)},\overline{\underset{\sim}{x}},a)dx_1 \ldots dx_p. \tag{3.1.9}$$

In order to evaluate Integral (3.1.9) more conveniently, let

$$Z_i = \left[\frac{n(n-p-1)}{(n-1)a_{ii}}\right]^{\frac{1}{2}} \frac{X_i^{(1)} - \bar{X}_i}{\left[1 - \left(\frac{n}{n-1}\right)\left(\underset{\sim}{X}^{(1)}-\bar{\underset{\sim}{X}}\right)' A^{-1}\left(\underset{\sim}{X}^{(1)}-\bar{\underset{\sim}{X}}\right)\right]^{\frac{1}{2}}}, \quad i = 1, 2, \ldots, p$$

where a_{ii} are the diagonal elements of A.

The Jacobian (A.12, Appendix) is

$$\frac{\partial\left(x_1^{(1)}, \ldots, x_p^{(1)}\right)}{\partial(Z_1, \ldots, Z_p)}$$

$$= \left(\frac{n-1}{n}\right)^{\frac{p}{2}} \frac{\prod_{i=1}^{p}\sqrt{a_{ii}}}{(n-p-1)^{\frac{p}{2}}} \left[1 + \frac{(\underset{\sim}{aZ})' A^{-1} (\underset{\sim}{aZ})}{n-p-1}\right]^{-\frac{1}{2}(p+2)}$$

where

$$\underset{\sim}{a}Z' = \left(\sqrt{a_{11}}Z_1, \quad \sqrt{a_{22}}Z_2, \ \ldots, \ \sqrt{a_{pp}}Z_p\right)$$

and where (A.12, Appendix)

$$1 - \frac{n}{n-1}\left(\underset{\sim}{X}^{(1)}-\bar{\underset{\sim}{X}}\right)' A^{-1}\left(\underset{\sim}{X}^{(1)}-\bar{\underset{\sim}{X}}\right) = \left(1 + \frac{(a\underset{\sim}{Z})' A^{-1} (a\underset{\sim}{Z})}{n-p-1}\right)^{-1}.$$

Integral (3.1.9) becomes

$F(Z_1,\ldots,Z_p)$

$$= \int_{-}^{Z_p} \cdots \int_{-}^{Z_1} \frac{\left(\frac{n-1}{2}\right)}{\left[\pi(n-p-1)\right]^{\frac{1}{2}} \Gamma\left(\frac{n-p-1}{2}\right)|R|^{\frac{1}{2}}} \left(1 + \frac{\underset{\sim}{Z}' R^{-1} \underset{\sim}{Z}}{n-p-1}\right)^{-\frac{1}{2}(n-1)} dZ_1 \ldots dZ_p$$

$$\underset{\sim}{Z} \in R^p$$

where R is the sample correlation matrix and where

$$Z_i(h_1,\ldots,h_p) = \left[\frac{n(n-p-1)}{(n-1)a_{ii}}\right]^{\frac{1}{2}} \frac{h_i-\bar{x}_i}{\left[1-\left(\frac{n}{n-1}\right)(\underset{\sim}{h}-\underset{\sim}{\bar{x}})A^{-1}(\underset{\sim}{h}-\underset{\sim}{\bar{x}})\right]^{\frac{1}{2}}}, \quad i = 1, \ldots, p.$$

Equation (3.1.10) is the cumulative distribution of a p-variate t variate with n-p-1 degrees of freedom [8].

CHAPTER IV

SUMMARY

In the study of a bivariate normal population for which an estimate of the probability $F(h,k)$ is desired, one common approach has been to replace the parameter (or parameters) appearing in the distribution function by estimates of the parameters. The parameters are usually either maximum likelihood estimates or unbiased estimates. The use of maximum likelihood estimates of the parameters will give a maximum likelihood estimate of the distribution function. However, if unbiased estimates of $F(h,k)$ are desired, these cannot necessarily be obtained by the use of unbiased parameter estimates. Just to illustrate this point, suppose that one wishes to find an estimate of the probability $F(h|\mu,\sigma^2)$ where F is the cumulative normal distribution with unknown mean and known variance. One common approach has been to find $\overline{x}$ from a random sample. Then the use of a cumulative normal distribution by substituting $\overline{x}$ for μ gives $F_z\left[(h-\overline{x})/\sigma\right]$ as an estimate of $F(h|\mu,\sigma^2)$. But the unbiased estimate of $F(h|\mu,\sigma^2)$ as given in Chapter II is $F_z\left[\sqrt{n}(h-\overline{x})/\sigma\sqrt{n-1}\right]$. Hence, $F_z\left[(h-\overline{x})/\sigma\right]$ is biased for $F(h|\mu,\sigma^2)$. This thesis offers the problem a solution which is not only direct but also the best in the sense of minimum variance and unbiasedness.

Even though only six cases of the parameters, in terms of their being known or unknown, of a bivariate normal density have been presented, it has been found that for many of the unpresented cases there exist no complete sufficient statistics. One such case, for instance, is when the correlation and means are known and the variances are unknown. Hence the Rao-Blackwell and Lehmann-Scheffé theorems do not apply.

In the Case 2.1 where all the parameters are unknown, the best estimate of the desired probability has been found for all sample sizes except n = 3. For n = 3,

$$T = \left\{ \sum_{i=1}^{3} X_i, \quad \sum_{i=1}^{3} Y_i, \quad \sum_{i=1}^{3} X_i^2, \quad \sum_{i=1}^{3} Y_i^2, \quad \sum_{i=1}^{3} X_i Y_i \right\}$$

is complete sufficient for the five parameters. The best estimate of the probability is given by

$$P(X_1 < h, \; Y_1 < k \mid T).$$

However, this writer has not been able to express this as an explicit function of h, k, and T. This is typical of several of the other cases in that the solution for very small sample sizes is more difficult than for larger sample sizes. The small sample size problems are of no practical significance as far as can be seen, but present interesting problems.

There still remain some cases that could be of interest. For instance, what would be the best estimate of the desired probability for the case when some subsets of the mean vector are equal but unknown in conjunction with a known or an unknown variance-covariance matrix? Perhaps of foremost interest is the problem of finding the exact variance of the estimators. If the problem is formidable, an expression of the lower bound or even an upper bound of the variance of the estimators may be manageable.

A SELECTED BIBLIOGRAPHY

[1] Folks, J. L., D. A. Pierce, and C. Stewart. "Estimating the Fraction of Acceptable Product." Technometrics, Vol. 7 (1965), 43-50.

[2] Hogg, R. V., and A. T. Craig. Introduction to Mathematical Statistics. New York: The MacMillan Company, 1970.

[3] Kempthorne, O., and J. L. Folks. Probability, Statistics, and Data Analysis. Ames, Iowa: Iowa State University Press, 1971.

[4] Basu, A. P. "Estimates of Reliability for Some Distributions Useful in Life Testing." Technometrics, Vol. 6 (1964), 215-219.

[5] Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, 1958.

[6] Dunnett, C. W., and M. Sobel. "A Bivariate Generalization of Student's t-Distribution with Tables for Certain Special Cases." Biometrics, Vol. 41 (1954), 153-169.

[7] National Bureau of Standards. Tables of the Bivariate Normal Distribution Function and Related Functions. Washington, D. C.: U. S. Department of Commerce, 1959.

[8] Johnson, N. L., and S. Kotz. Continuous Multivariate Distributions. New York: John Wiley and Sons, 1972.

[9] Wasan, M. T. Parametric Estimation. New York: McGraw-Hill, 1970.

[10] Lehmann, E. L., and H. Scheffé. "Completeness, Similar Regions, and Unbiased Estimation." Sankyā, Ser. A., Pt. 4, Vol. 10 (1950), 305-340.

APPENDIX

A.1. Let (X_1,Y_1), (X_2,Y_2), ..., (X_n,Y_n) be a random sample of size n from a bivariate normal distribution with means μ_x and μ_y, positive variance σ_x^2 and σ_y^2, and correlation coefficient ρ, where $0 < |\rho| < 1$. Show that

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i, \quad A_{11} = \sum_{i=1}^{n} (X_i-\overline{X})^2,$$

and

$$A_{12} = \sum_{i=1}^{n} (X_i-\overline{X})(Y_i-\overline{Y}), \quad A_{22} = \sum_{i=1}^{n} (Y_i-\overline{Y})^2$$

are jointly complete sufficient for the five parameters.

Proof: The joint density of (X_1,Y_1), (X_2,Y_2), ..., (X_n,Y_n) is

$$f((x_1,y_1),\ (x_2,y_2),\ \ldots,\ (x_n,y_n)\,|\,\mu_x,\ \mu_y,\ \sigma_x,\ \sigma_y,\ \rho) \tag{A.1.1}$$

$$= \frac{1}{(2\pi)^n\sigma_x^n\sigma_y^n(1-\rho^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2(1-\rho^2)}\sum_{1}^{n}\left[\left(\frac{x_i-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x_i-\mu_x}{\sigma_x}\right)\left(\frac{y_i-\mu_y}{\sigma_y}\right) + \left(\frac{y_i-\mu_y}{\sigma_y}\right)^2\right]\right.$$

$$= \frac{1}{(2\pi)^n \sigma_x^n \sigma_y^n (1-\rho^2)^{\frac{n}{2}}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \sum_1^n \left[\frac{1}{\sigma_x^2} x_i^2 - \frac{2\rho}{\sigma_x \sigma_y} x_i y_i \right.\right.$$

$$\left.\left. + \frac{1}{\sigma_y^2} y_i^2 + \left(\frac{2\rho\mu_y}{\sigma_x\sigma_y} - \frac{2\mu_x}{\sigma_x^2} \right) x_i + \left(\frac{2\rho\mu_x}{\sigma_x\sigma_y} - \frac{2\mu_y}{\sigma_y^2} \right) y_i + \frac{\mu_x^2}{\sigma_x^2} - \frac{2\rho\mu_x\mu_y}{\sigma_x\sigma_y} + \frac{\mu_y^2}{\sigma_y^2} \right]$$

Equation (A.1.1) can be written in the form

$$f(x,y|\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$$

$$= \exp \left[\sum_{j=1}^{5} \phi_j(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho) T_j(x,y) + h(x,y) + C(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho) \right]$$

with

$$\phi_1 = \frac{1}{2(1-\rho^2)\sigma_x^2}, \quad \phi_2 = \frac{1}{2(1-\rho^2)\sigma_y^2}, \quad \phi_3 = \frac{2\rho}{2(1-\rho^2)\sigma_x\sigma_y},$$

$$\phi_4 = \frac{1}{2(1-\rho^2)} \left(\frac{2\rho\mu_y}{\sigma_x\sigma_y} - \frac{2\mu_x}{\sigma_x^2} \right), \quad \phi_5 = \frac{1}{2(1-\rho^2)} \left(\frac{2\rho\mu_x}{\sigma_x\sigma_y} - \frac{2\mu_y}{\sigma_y^2} \right), \tag{A.1.2}$$

and

$$T_1 = \sum_{i=1}^{n} X_i^2, \quad T_2 = \sum_{i=1}^{n} Y_i^2, \quad T_3 = \sum_{i=1}^{n} X_i Y_i,$$

$$T_4 = \sum_{i=1}^{n} X_i, \quad T_5 = \sum_{i=1}^{n} Y_i \tag{A.1.3}$$

and

$$C(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho) = - \ln\left[(2\pi)^n\sigma_x^n\sigma_y^n(1-\rho^2)^{\frac{n}{2}}\right],$$

and

$$h(x,y) = 0.$$

The parameter space

$$\Omega = \{-\infty < \mu_x,\mu_y < \infty, \sigma_x^2,\sigma_y^2 > 0,\ 0 < |\rho| > 1\}$$

contains a five-dimensional rectangle defined by Equation (A.1.2), hence, by Lehmann and Scheffé [9], the statistics defined by Equation (A.1.3) are jointly complete sufficient for the parameters. The transformation between Equation (A.1.3) and $\overline{X}$, $\overline{Y}$, A_{11}, A_{12}, A_{22} is one to one. Hence $\overline{X}$, $\overline{Y}$, A_{11}, A_{12}, A_{22} are jointly complete sufficient for μ_x, μ_y, σ_x^2, σ_y^2 and ρ.

A.2. Let (X_1,Y_1), (X_2,Y_2), ..., (X_n,Y_n) be a random sample from a bivariate distribution. Let

$$\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}, \quad \overline{Y} = \sum_{i=1}^{n} \frac{Y_i}{n},$$

$$A_{11} = \sum_{i=1}^{n} (X_i-\overline{X})^2, \quad A_{22} = \sum_{i=1}^{n} (Y_i-\overline{Y})^2,$$

$$A_{12} = \sum_{i=1}^{n} (X_i-\overline{X})(Y_i-\overline{Y}),$$

$$\overline{X}^* = \sum_{i=2}^{n} \frac{X_i}{n-1}, \quad \overline{Y}^* = \sum_{i=2}^{n} \frac{Y_i}{n-1},$$

$$A_{11}^* = \sum_{i=2}^{n} (X_i-\overline{X}^*)^2, \quad A_{22}^* = \sum_{i=2}^{n} (Y_i-\overline{Y}^*)^2,$$

$$A_{12}^* = \sum_{i=2}^{n} (X_i-\overline{X}^*)(Y_i-\overline{Y}^*).$$

Show that

(a) $A_{11} = A_{11}^* + \frac{n-1}{n} (X_1-\overline{X}^*)^2,$

(b) $A_{12} = A_{12}^* + \frac{n-1}{n} (X_1-\overline{X}^*)(Y_1-\overline{Y}^*).$

Proof:

(a) $$A_{11} = \sum_{i=1}^{n} (X_i-\overline{X})^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2$$

$$= \sum_{i=1}^{n} X_i^2 - \frac{[(n-1)\overline{X}^* + X_1]^2}{n}$$

$$= \sum_{i=2}^{n} X_i^2 + X_1^2 - \frac{(n-1)^2\overline{X}^{*2} + 2(n-1)\overline{X}^*X_1 + X_1^2}{n}$$

$$= \sum_{i=2}^{n} X_i^2 - (n-1)\overline{X}^{*2} + \frac{n-1}{n} (X_1-\overline{X}^*)^2$$

$$= A_{11}^* + \frac{n-1}{n} (X_1-\overline{X}^*)^2 .$$

(b) $$A_{12} = \sum_{i=1}^{n} (X_i-\overline{X})(Y_i-\overline{Y}) = \sum_{i=1}^{n} X_iY_i - nXY$$

$$= \sum_{i=2}^{n} X_iY_i + X_1Y_1 - \frac{[(n-1)\overline{X}^*+X_1]\ [(n-1)\overline{Y}^*+Y_1]}{n}$$

$$= \sum_{i=2}^{n} X_iY_i - \frac{(n-1)X_1Y_1-(n-1)(\overline{X}^*Y_1+\overline{Y}^*X_1)+(n-1)\overline{X}^*\overline{Y}^*-n(n-1)\overline{X}^*\overline{Y}^*}{n}$$

$$= \sum_{i=2}^{n} X_iY_i - (n-1)\overline{X}^*\overline{Y}^* + \frac{n-1}{n}\left[X_1Y_1 - \overline{X}^*Y_1 - \overline{Y}^*X_1 + \overline{X}^*\overline{Y}^*\right]$$

$$= A_{12}^* + \frac{n-1}{n}(X_1-\overline{X}^*)(Y_1-\overline{Y}^*).$$

A.3. Show that

$$f_{X_1,Y_1\ T}(x_1,y_1|\overline{x},\overline{y},a_{11},a_{12},a_{22}) = \frac{(n-1)\Gamma\left(\frac{n-1}{2}\right)}{n\pi\Gamma\left(\frac{n-3}{2}\right)\ \left(a_{11}a_{22}-a_{12}^2\right)^{\frac{1}{2}}}\ \cdot$$

$$\left[1 - \frac{n}{n-1}\ \frac{(\overline{x}-x_1)^2a_{22} - 2(\overline{x}-x_1)(\overline{y}-y_1)a_{12} + (\overline{y}-y_1)a_{11}}{a_{11}a_{22} - a_{12}^2}\right]^{\frac{1}{2}(n-5)}.$$

Proof:

$$f_{X_1,Y_1\ T}(x_1,y_1|\overline{x},\overline{y},a_{11},a_{12},a_{22}) = \frac{f_{X_1,Y_1,T}(x_1,y_1,\overline{x},\overline{y},a_{11},a_{12},a_{22})}{f_T(\overline{x},\overline{y},a_{11},a_{12},a_{22})} =$$

$$\frac{(n-1)\left\{\left[a_{11}-\frac{n}{n-1}(\overline{x}-x_1)^2\right]\left[a_{22}-\frac{n}{n-1}(\overline{y}-y_1)^2\right]-\left[a_{12}-\frac{n}{n-1}(\overline{x}-x_1)(\overline{y}-y_1)^2\right]\right\}^{\frac{1}{2}(n-5)}}{2^n\pi^{\frac{5}{2}}\sigma_x^n\sigma_y^n(1-\rho^2)^{\frac{n}{2}}\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n-3}{2}\right)}\cdot$$

$$\exp -\frac{1}{2(1-\rho^2)}\cdot\left\{\left[\left(\frac{x_1-\mu_x}{\sigma_x}\right)^2-2\rho\left(\frac{x_1-\mu_x}{\sigma_x}\right)\left(\frac{y_1-\mu_y}{\sigma_y}\right)+\left(\frac{y_1-\mu_y}{\sigma_y}\right)^2\right]+\right.$$

$$(n-1)\left[\left(\frac{\frac{n\overline{x}-x_1}{n-1}-\mu_x}{\sigma_x}\right)^2-2\rho\left(\frac{\frac{n\overline{x}-x_1}{n-1}-\mu_x}{\sigma_x}\right)\left(\frac{\frac{n\overline{y}-y_1}{n-1}-\mu_y}{\sigma_y}\right)+\left(\frac{\frac{n\overline{y}-y_1}{n-1}-\mu_y}{\sigma_y}\right)^2\right]+$$

$$\left.\frac{a_{11}-\frac{n}{n-1}(\overline{x}-x_1)^2}{\sigma_x^2}-2\rho\,\frac{a_{12}-\frac{n}{n-1}(\overline{x}-x_1)(\overline{y}-y_1)}{\sigma_x\sigma_y}+\frac{a_{22}-\frac{n}{n-1}(\overline{y}-y_1)^2}{\sigma_y^2}\right\}\Bigg/$$

$$\frac{n(a_{11}a_{22}-a_{12}^2)^{\frac{1}{2}(n-4)}}{2^n\pi^{\frac{3}{2}}\sigma_x^n\sigma_y^n(1-\rho^2)^{\frac{n}{2}}\;\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}\exp\left\{-\frac{n}{2(1-\rho^2)}\left[\left(\frac{\overline{x}-\mu_x}{\sigma_x}\right)^2-2\rho\left(\frac{\overline{x}-\mu_x}{\sigma_x}\right)\left(\frac{\overline{y}-\mu_y}{\sigma_y}\right)\right.\right.$$

$$\left.\left.+\left(\frac{\overline{y}-\mu_y}{\sigma_y}\right)^2\right]+\frac{a_{11}}{\sigma_x^2}-2\rho\,\frac{a_{12}}{\sigma_x\sigma_y}+\frac{a_{22}}{\sigma_y^2}\right\}.$$

The constant of the quotient is

$$\frac{(n-1)2^n\pi^{\frac{3}{2}}\sigma_x^n\sigma_y^n(1-\rho^2)^{\frac{n}{2}}\;\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)}{n2^n\pi^{\frac{5}{2}}\sigma_x^n\sigma_y^n(1-\rho^2)^{\frac{n}{2}}\;\Gamma\left(\frac{n-2}{2}\right)\Gamma\left(\frac{n-3}{2}\right)}=\frac{(n-1)\Gamma\left(\frac{n-1}{2}\right)}{n\pi^{\frac{1}{2}}\Gamma\left(\frac{n-3}{2}\right)}. \tag{A.3.1}$$

Since

$$\left(\frac{x_1-\mu_x}{\sigma_x}\right)^2 + (n-1)\left(\frac{\frac{n\bar{x}-x_1}{n-1} - \mu_x}{\sigma_x}\right)^2 - n\left(\frac{\bar{x}-\mu_x}{\sigma_x}\right)^2 = \frac{n}{n-1}\left(\frac{\bar{x}-x_1}{\sigma_x}\right)^2,$$

$$\left(\frac{y_1-\mu_y}{\sigma_y}\right)^2 + (n-1)\left(\frac{\frac{n\bar{y}-y_1}{n-1} - \mu_y}{\sigma_y}\right)^2 - n\left(\frac{\bar{y}-\mu_y}{\sigma_y}\right)^2 = \frac{n}{n-1}\left(\frac{\bar{y}-y_1}{\sigma_y}\right)^2,$$

$$-\frac{(x_1-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} - (n-1)\frac{\left(\frac{n\bar{x}-x_1}{n-1} - \mu_x\right)\left(\frac{n\bar{y}-y_1}{n-1} - \mu_y\right)}{\sigma_x\sigma_y} + \frac{n(\bar{x}-\mu_x)(\bar{y}-\mu_y)}{\sigma_x\sigma_y}$$

$$= -\frac{n}{n-1}\frac{(\bar{x}-x_1)(\bar{y}-y_1)}{\sigma_x\sigma_y},$$

the exponents reduce to zero. Thus since $\exp\{0\} = 1$, the expression takes the form

$$\frac{\left\{\left[a_{11} - \frac{n}{n-1}(\bar{x}-x_1)^2\right]\left[a_{22} - \frac{n}{n-1}(\bar{y}-y_1)^2\right] - \left[a_{12} - \frac{n}{n-1}(\bar{x}-x_1)(\bar{y}-y_1)\right]^2\right\}^{\frac{1}{2}(n-5)}}{\left(a_{11}a_{22} - a_{12}^2\right)^{\frac{1}{2}(n-4)}}$$

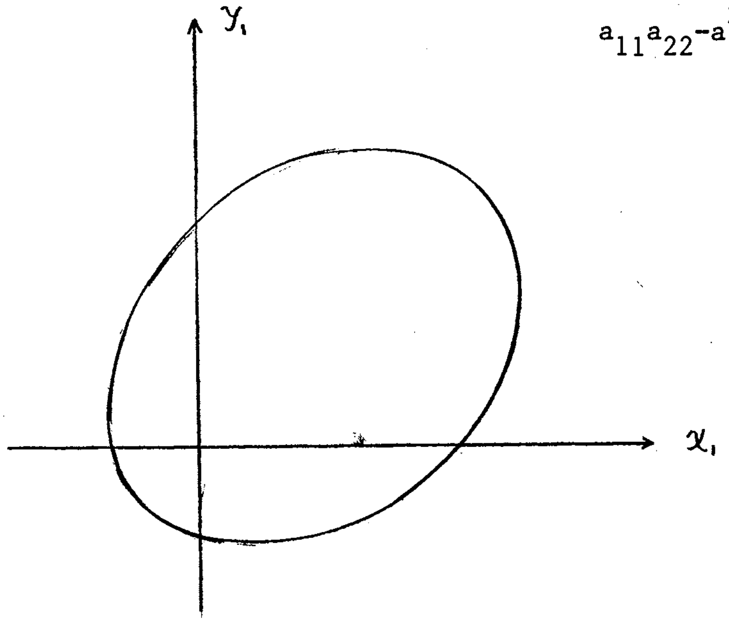
$$= \frac{\left\{a_{11}a_{22}-a_{12}^2 - \frac{n}{n-1}\left[(\bar{x}-x_1)^2a_{22} - 2(\bar{x}-x_1)(\bar{y}-y_1)a_{12} + (\bar{y}-y_1)^2a_{11}\right]\right\}^{\frac{1}{2}(n-5)}}{\left(a_{11}a_{22} - a_{12}^2\right)^{\frac{1}{2}(n-4)}} \tag{A.3.2}$$

$$= \frac{1}{\left(a_{11}a_{22} - a_{12}^2\right)^{\frac{1}{2}}}\left[1 - \frac{n}{n-1}\,\frac{(\bar{x}-x_1)^2a_{22}-2(\bar{x}-x_1)(\bar{y}-y_1)a_{12}+(\bar{y}-y_1)^2a_{11}}{a_{11}a_{22} - a_{12}^2}\right]^{\frac{1}{2}(n-5)}$$

The proof is completed by multiplying Equations (A.3.1) and (A.3.2) together.

A.4. The domain of Equation (2.1.7)

$$\frac{a_{22}(x_1-\bar{x})^2-2a_{12}(x_1-\bar{x})(y_1-\bar{y})+a_{11}(y_1-\bar{y})^2}{a_{11}a_{22}-a_{12}^2} < \frac{n-1}{n}$$



A.5. Show that under transformation (2.1.9), Estimate (2.1.8) becomes Estimate (2.1.11).

Proof: Given that

$$f_{X_1,Y_1|T}(x_1,y_1\ \bar{x},\bar{y},a_{11},a_{12},a_{22})$$

$$= \frac{(n-1)\Gamma\left(\frac{n-1}{2}\right)}{n\pi\Gamma\left(\frac{n-3}{2}\right)|A|^{\frac{1}{2}}}\left[1-\frac{n}{n-1}\,\frac{a_{22}(x_1-\bar{x})^2-2a_{12}(x_1-\bar{x})(y_1-\bar{y})+a_{11}(y_1-\bar{y})^2}{|A|}\right]^{\frac{1}{2}(n-5)} . \quad \text{(A.5.1)}$$

Let

$$T_1 = \frac{X_1-\overline{X}}{\sqrt{\left(\frac{n-1}{n}\right)a_{11}}}, \quad T_2 = \frac{}{\sqrt{\left(\frac{n-1}{n}\right)a_{22}}}, \; n > 1.$$

Equation (A.5.1) in t_1 and t_2 takes the form

$$f(t_1,t_2) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\pi\Gamma\left(\frac{n-3}{2}\right)(1-r^2)^{\frac{1}{2}}}\,(1 - \underset{\sim}{t}'R^{-1}\underset{\sim}{t})^{\frac{1}{2}(n-5)} = \frac{\Gamma\left(\frac{n-1}{2}\right)\sqrt{a_{11}a_{22}}}{\pi\Gamma\left(\frac{n-3}{2}\right)|A|^{\frac{1}{2}}} \; . \tag{A.5.2}$$

$$\left(1 - \frac{a_{11}a_{22}t_1^2 - 2a_{12}\sqrt{a_{11}a_{22}}t_1t_2 + a_{11}a_{22}t_2^2}{|A|}\right)^{\frac{1}{2}(n-5)}$$

where

$$r = \frac{a_{12}}{\sqrt{a_{11}a_{22}}}, \quad \underset{\sim}{t}' = (t_1,t_2), \quad R^{-1} = \begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix} \Big/ (1-r^2).$$

Let

$$\underset{\sim}{t}'R^{-1}\underset{\sim}{t} = Q_t,$$

$$\frac{Z_1^2-2rZ_1Z_2+Z_2^2}{1-r^2} = Q_Z.$$

Set

$$1 - Q_t = \frac{1}{1 + \frac{Q_Z}{n-3}},$$

then

$$Q_t = \frac{Q_Z}{(n-3) + Q_Z} \ . \qquad (A.5.3)$$

Also,

$$Q_Z = \frac{(n-3)Q_t}{1 - Q_t} \ . \qquad (A.5.4)$$

Express t in terms of Z by the following equation

$$(t_1,t_2)R^{-1}\begin{pmatrix} t_1 \\ \\ t_2 \end{pmatrix} = \frac{1}{(n-3) + Q_Z} (Z_1,Z_2)R^{-1}\begin{pmatrix} Z_1 \\ \\ Z_2 \end{pmatrix} \qquad (A.5.5)$$

$$= \left(\frac{Z_1}{[(n-3)+Q_Z]^{\frac{1}{2}}}, \frac{Z_2}{[(n-3)+Q_Z]^{\frac{1}{2}}}\right) R^{-1} \begin{pmatrix} \dfrac{Z_1}{[(n-3)+Q_Z]^{\frac{1}{2}}} \\ \\ \dfrac{Z_2}{[(n-3)+Q_Z]^{\frac{1}{2}}} \end{pmatrix} .$$

This suggests the transformation

$$t_1 = \frac{Z_1}{[(n-3)+Q_Z]^{\frac{1}{2}}} , \qquad t_2 = \frac{Z_2}{[(n-3)+Q_Z]^{\frac{1}{2}}} \ .$$

Equation (A.5.5) also gives

$$Z_1 = \frac{(n-3)^{\frac{1}{2}} t_1}{(1-Q_t)^{\frac{1}{2}}} , \qquad Z_2 = \frac{(n-3)^{\frac{1}{2}} t_2}{(1-Q_t)^{\frac{1}{2}}} \ .$$

The Jacobian is

$$\frac{\partial(t_1,t_2)}{\partial(Z_1,Z_2)} = \frac{1}{n-3}\left(1+\frac{1}{n-3}Q_Z\right)^{-2}.$$

Equation (A.5.2) in terms of Z_1 and Z_2 takes the form

$$f(Z_1,Z_2) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{|R|^{\frac{1}{2}}(n-3)\pi\Gamma\left(\frac{n-3}{2}\right)}\left(1+\frac{1}{n-3}\underset{\sim}{Z}'R^{-1}\underset{\sim}{Z}\right)^{-\frac{1}{2}(n-1)}, \quad \underset{\sim}{Z}\in R^2. \qquad \text{(A.5.6)}$$

where $|R| = (1-r^2)$. Equation (A.5.6) is in the form

$$f(Z_1,Z_2) = \frac{\Gamma\left[\frac{1}{2}(v+p)\right]}{(\pi v)^{\frac{1}{2}p}\Gamma\left(\frac{v}{2}\right)|R|^{\frac{1}{2}}}(1+v^{-1}\underset{\sim}{Z}'R^{-1}\underset{\sim}{Z})^{-\frac{1}{2}(v+p)}$$

with $v = n - 3$ and $p = 2$. It is then a bivariate t density with $n - 3$ degrees of freedom.

A.6. The proof is the same as that given in A.1 with n being one. This fact shows that completeness does not depend on sample size.

A.7. Given $T = t$, that is, $X_1 + X_2 = t_1$, $X_1^2 + X_2^2 = t_2$, $Y_1 + Y_2 = t_3$, $Y_1^2 + Y_2^2 = t_4$, $X_1Y_1 + X_2Y_2 = t_5$, the solution set in the $X_1Y_1X_2Y_2$-space consists of two points. Let them be denoted by $(\ell_1, m_1, \ell_2, m_2)$ and $(\ell_2, m_2, \ell_1, m_1)$; and let $\ell_1 < \ell_2$, $m_1 < m_2$. Then the probability mass function at $(\ell_1, m_1, \ell_2, m_2)$ is derived heuristically in the following.

$$p\left(x_1,y_1,x_2,y_2 \,\middle|\, (\ell_1,m_1,\ell_2,m_2),(\ell_2,m_2,\ell_1,m_1)\right)$$

$$= \lim_{\substack{h_1\to 0,\ k_1\to 0 \\ h_2\to 0,\ k_2\to 0}}$$

$$\frac{p(\ell_1-h_1<x_1<\ell_1+h_1, m_1-k_1<y_1<m_1+k_1, \ell_2-h_2<x_2<\ell_2\ +h_2, m_2-k_2<y_2<m_2+k_2)}{p(\ell_1-h_1<x_1<\ell_1<h_1,\ldots,m_2-k_2<y_2<m_2+k_2)+p(\ell_2-h_1<x_1<\ell_2+h_1,\ldots,m_1-k_2<y_2<m_1+k_2)}$$

$$= \lim_{\substack{h_1\to 0,\ k_1\to 0 \\ h_2\to 0,\ k_2\to 0}} \frac{h_1k_1f(\ell_1,m_1)h_2k_2f(\ell_2,m_2)}{h_1k_1f(\ell_1,m_1)h_2k_2f(\ell_2,m_2)+h_1k_1f(\ell_2,m_2)h_2k_2f(\ell_1,m_1)}$$

$$= \lim_{\substack{h_1\to 0,\ k_1\to 0 \\ h_2\to 0,\ k_2\to 0}} \frac{h_1k_1h_2k_2f(\ell_1,m_1)f(\ell_2,m_2)}{2h_1k_1h_2k_2f(\ell_1,m_1)f(\ell_2,m_2)}$$

$$= \frac{1}{2}$$

where h_1, k_1, h_2, and k_2 are small positive quantities. With this fact the result of the case for $n = 2$ is easily obtained.

A.8. Equation (2.4.1) is in the regular exponential class with

$$\phi_x = \frac{1}{\sigma_x^2}, \quad \phi_2 = \frac{1}{\sigma_y^2}, \quad T_1 = \sum_{i=1}^{n}(X_i-\mu_x)^2, \quad T_2 = \sum_{i=1}^{n}(Y_i-\mu_y)^2.$$

The parameter space $\Omega = \{(0,\infty)\ X\ (0,\infty)\}$ contains a two-dimensional rectangle. So $\Sigma(X_i-\mu_x)^2$ and $\Sigma(Y_i-\mu_y)^2$ are complete sufficient for σ_x^2 and σ_y^2.

A.9. If $\underset{\sim}{X} \sim n_p(\underset{\sim}{x}|\underset{\sim}{\mu},\Sigma)$, then $\overline{X}$ and A are complete sufficient for $\underset{\sim}{\mu}$ and Σ.

Proof: Let $\{X^{(1)}, \ldots, X^{(n)}\}$ and $\{X^{(1)^o}, \ldots, X^{(n)^o}\}$ be two random samples of size n each. Then

$$\frac{\left(f\ \underset{\sim}{x}^{(1)},\ \ldots,\ \underset{\sim}{x}^{(n)}\right)}{f\left(\underset{\sim}{x}^{(1)^{o}},\ \ldots,\ \underset{\sim}{x}^{(n)^{o}}\right)}$$

$$= \frac{\dfrac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left\{-\frac{1}{2}\sum_{k=1}^{n}\left[(\underset{\sim}{x}^{(k)}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{x}^{(k)}-\underset{\sim}{\mu})\right]\right\}}{\dfrac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left\{-\frac{1}{2}\sum_{k=1}^{n}\left[(\underset{\sim}{x}^{(k)^{o}}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{x}^{(k)^{o}}-\underset{\sim}{\mu})\right]\right\}}$$

$$= \frac{\exp\left\{-\frac{1}{2}\left[\mathrm{tr}\Sigma^{-1}A+n(\underset{\sim}{\bar{x}}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{\bar{x}}-\underset{\sim}{\mu})\right]\right\}}{\exp\left\{-\frac{1}{2}\left[\mathrm{tr}\Sigma^{-1}A^{o}+n(\underset{\sim}{\bar{x}}^{o}-\underset{\sim}{\mu})\Sigma^{-1}(\underset{\sim}{\bar{x}}-\underset{\sim}{\mu})\right]\right\}}$$

$$= \exp\left\{-\frac{1}{2}\left[\mathrm{tr}\Sigma^{-1}A-\mathrm{tr}\ A^{o}+(\underset{\sim}{\bar{x}}-\underset{\sim}{\bar{x}}^{o})'\Sigma^{-1}(\underset{\sim}{\bar{x}}-\underset{\sim}{\bar{x}}^{o})\right]\right\}.$$

This ratio is free of Σ^{-1} if $A = A^{o}$ and $\underset{\sim}{\bar{x}} = \underset{\sim}{\bar{x}}^{o}$. Then by Lehmann and Sheffé [10], $T = \{\bar{X}, A\}$ is minimal sufficient for $\underset{\sim}{\mu}$ and Σ. Now

$$f_T(\underset{\sim}{\bar{x}},A) = \frac{n^{p/2}}{(2\pi)^{p/2}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{n}{2}(\underset{\sim}{\bar{x}}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{\bar{x}}-\underset{\sim}{\mu})\right]\cdot$$

$$\frac{|A|^{\frac{1}{2}(n-2-p)} \exp\left(-\frac{1}{2}\mathrm{tr}\Sigma^{-1}A\right)}{2^{\frac{1}{2}(n-1)p}\ \pi^{\frac{1}{4}p(p-1)}\ |\Sigma|^{\frac{1}{2}(n-1)} \prod_{i=1}^{p}\Gamma\left[\frac{1}{2}(n-i)\right]}.$$

Let Σ^{-1} be denoted by Ψ.

$$L = \log f_T(\underset{\sim}{\bar{x}},A)$$

$$= \frac{p}{2}\log n - \frac{p}{2}\log 2\pi + \frac{1}{2}(n-2-p)\log|A| - \frac{1}{2}(n-1)p\log 2 -$$

$$- \frac{1}{4}p(p-1)\ \log\ \pi + \frac{n-1}{2}\ \log|\Psi| + \frac{1}{2}\ \mathrm{tr}\Psi A - \frac{1}{2}n(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})'\Psi(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})$$

$$= C + \frac{1}{2}\ n\ \log|\Psi| - \frac{1}{2}\ \mathrm{tr}\Psi A - \frac{1}{2}\ n(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})'\Psi(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})$$

where C denotes the sum of the constant terms. This L has exactly the same expression as given by Anderson [5]. The values of μ and Ψ that maximize L are $\bar{\underset{\sim}{X}}$ and nA^{-1} respectively. But the relation between $\{\bar{\underset{\sim}{X}}, A\}$ and $\{\bar{\underset{\sim}{X}}, nA^{-1}\}$ is one-to-one. Hence by Kempthorne and Folks [3], $T = \{\bar{\underset{\sim}{X}}, A\}$ is F-sufficient for $\underset{\sim}{\mu}$ and Σ. This implies that $n_p(\underset{\sim}{x}|\underset{\sim}{\mu},\Sigma)$ is of the FKPD form and $T = \{\bar{\underset{\sim}{X}}, A\}$ is complete sufficient for $\underset{\sim}{\mu}$ and Σ.

A.10. Show that

$$\left|\frac{\partial \underset{\sim}{x}^{(1)}, \bar{\underset{\sim}{x}}^{*}, A^{*}}{\partial \underset{\sim}{x}^{(1)}, \bar{\underset{\sim}{x}}, A}\right| = \left(\frac{n}{n-1}\right)^{p}$$

where

$$\underset{\sim}{X}^{(1)} = \underset{\sim}{X}^{(1)}, \quad \bar{\underset{\sim}{X}}^{*} = \frac{n}{n-1}\,\bar{\underset{\sim}{X}} - \frac{n}{n-1}\,\underset{\sim}{X}^{(1)}$$

$$A^{*} = A - \frac{n}{n-1}\left(\bar{\underset{\sim}{X}} - \underset{\sim}{X}^{(1)}\right)\left(\bar{\underset{\sim}{X}} - \underset{\sim}{X}^{(1)}\right)'$$

Proof:

$\partial(\underset{\sim}{x}^{(1)}, \bar{\underset{\sim}{x}}^{*}, a^{*})/\partial(\underset{\sim}{x}^{(1)}, \bar{\underset{\sim}{x}}, a)$ is the determinant of a matrix whose elements are skematically given by:

$$
\begin{array}{c|cccc|cccc|cccc|cccc}
 & x_1^{(1)} & \cdots & & x_p^{(1)} & \overline{x}_1 & \cdots & & \overline{x}_p & a_{11} & \cdots & & a_{pp} & a_{12} & \cdots & & a_{p-1,p} \\
\hline
x_1^{(1)} & 1 & & & 0 & & & & & & & & & & & & \\
\vdots & & \ddots & & & & 0 & & & & 0 & & & & 0 & & \\
x_p^{(1)} & 0 & & & 1 & & & & & & & & & & & & \\
\hline
\overline{x}_1^* & & & & & \frac{-n}{n-1} & & & & & & & & & & & \\
\vdots & & C_1 & & & & \ddots & & & & 0 & & & & 0 & & \\
\overline{x}_p^* & & & & & & & & \frac{-n}{n-1} & & & & & & & & \\
\hline
a_{11}^* & & & & & & & & & 1 & & & & & & & \\
\vdots & & & & & & & & & & \ddots & & & & 0 & & \\
a_{pp}^* & & & & & & & & & & & & 1 & & & & \\
a_{12}^* & & & & & & C_2 & & & & & & & 1 & & & \\
\vdots & & & & & & & & & & 0 & & & & \ddots & & \\
a_{p-1,p}^* & & & & & & & & & & & & & & & & 1
\end{array}
$$

where C_1 and C_2 are some non-zero submatrices which will not affect the determinant of this triangular matrix. The absolute value of the determinant of this matrix is seen to be

$$\left(\frac{n}{n-1}\right)^p, \; n > 1.$$

A.11. Derive Equation (3.1.8)

$$f_{X(1)|T}\left(\underset{\sim}{x}^{(1)} \mid \underset{\sim}{\overline{x}}, a\right) = \frac{f_{\underset{\sim}{X}(1),T}\left(\underset{\sim}{x}^{(1)}, \underset{\sim}{\overline{x}}, a\right)}{f_T(\underset{\sim}{\overline{x}}, a)} = \text{A} \times \text{B} \times \text{C}$$

where

$$\text{(A)} = \frac{\dfrac{n^p}{(n-1)^{p/2} 2^{np/2} \pi^{p(p+3)/4} |\Sigma|^{n/2} \prod_{i=1}^{p} \Gamma\left[\frac{1}{2}(n-1-i)\right]}}{\dfrac{n^{p/2}}{2^{np/2} \pi^{p(p+1)/4} |\Sigma|^{n/2} \prod_{i=1}^{p} \Gamma\left[\frac{1}{2}(n-i)\right]}}$$

$$= \left(\frac{n}{n-1}\right)^{p/2} \frac{\Gamma\left(\frac{n}{2}\right)}{\pi^{p/2} \Gamma\left(\frac{n-p}{2}\right)} ,$$

$$\text{(B)} = \frac{\left| A - \frac{n}{n-1} \left(\bar{\underset{\sim}{x}} - \underset{\sim}{x}^{(1)}\right) \left(\bar{\underset{\sim}{x}} - \underset{\sim}{x}^{(1)}\right)' \right|^{\frac{1}{2}(n-3-p)}}{|A|^{\frac{1}{2}(n-2-p)}}$$

$$= \frac{\left| |A| - \frac{n}{n-1} \sum_{i}^{p}\sum_{j}^{p} A_{ij} \left(x_i^{(1)} - \bar{x}_i\right) \left(x_j^{(1)} - \bar{x}_j\right) \right|^{\frac{1}{2}(n-3-p)}}{|A|^{\frac{1}{2}(n-2-p)}}$$

$$= \frac{|A|^{\frac{1}{2}(n-3-p)} \left| 1 - \frac{1}{n-1} \frac{\sum_{i}^{p}\sum_{j}^{p} A_{ij} \left(x_i^{(1)} - \bar{x}_i\right) \left(x_j^{(1)} - \bar{x}_j\right)}{|A|} \right|^{\frac{1}{2}(n-3-p)}}{|A|^{\frac{1}{2}(n-2-p)}}$$

$$= \frac{1}{|A|^{\frac{1}{2}}} \left[1 - \frac{n}{n-1} \left(\underset{\sim}{x}^{(1)} - \bar{\underset{\sim}{x}}\right)' A^{-1} \left(\underset{\sim}{x}^{(1)} - \bar{\underset{\sim}{x}}\right) \right]^{\frac{1}{2}(n-3-p)}$$

where A_{ij} are cofactors of a_{ij} and A^{-1} is the inverse of A.

$$\text{(C)} = \exp -\frac{1}{2}\left\{(n-1)\left[\frac{1}{n-1}\left(n\bar{\underset{\sim}{x}}-\underset{\sim}{x}^{(1)}\right)-\underset{\sim}{\mu}\right]'\Sigma^{-1}\left[\frac{1}{n-1}\left(n\bar{\underset{\sim}{x}}-\underset{\sim}{x}^{(1)}\right)-\underset{\sim}{\mu}\right] + \right.$$

$$\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}\right)'\Sigma^{-1}\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}\right)-n(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})'\Sigma^{-1}(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu}) +$$

$$\left. \operatorname{tr}\Sigma^{-1}\left[A - \frac{n}{n-1}\left(\bar{\underset{\sim}{x}}-\underset{\sim}{x}^{(1)}\right)\left(\bar{\underset{\sim}{x}}-\underset{\sim}{x}^{(1)}\right)'\right] - \operatorname{tr}\Sigma^{-1}A\right\}$$

$$= \exp -\frac{1}{2}\left\{\frac{1}{n-1}\left[n(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu}) + \left(\underset{\sim}{\mu}-\underset{\sim}{x}^{(1)}\right)\right]'\Sigma^{-1}\left[n(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu}) + \left(\underset{\sim}{\mu}-\underset{\sim}{x}^{(1)}\right)\right] + \right.$$

$$(n-1)\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}\right)'\Sigma^{-1}\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}\right) + n\left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right)'\Sigma^{-1}\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}\right) -$$

$$\left. n^2(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})'\Sigma^{-1}(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu}) + n(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})'\Sigma^{-1}(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})\Big]\right\}$$

$$= \exp\left\{-\frac{1}{2}\left[n\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}\right)'\Sigma^{-1}\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}\right) - n\left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right)'\Sigma^{-1}\left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right) + \right.\right.$$

$$\left.\left. n(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})'\Sigma^{-1}(\bar{\underset{\sim}{x}}-\underset{\sim}{\mu})\right]\right\}$$

$$= \exp\left\{-\frac{n}{2(n-1)}\left[\left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right)'\Sigma^{-1}\left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right) - \left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right)'\Sigma^{-1}\left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right)\right]\right.$$

$$= \exp\{0\} = 1.$$

The product of (A), (B), and (C) gives Equation (3.2.8).

A.12. Derive the integral (3.1.10).

Let

$$Z_i = \left(\frac{n(n-1-p)}{(n-1)a_{ii}}\right)^{\frac{1}{2}} \frac{x_i^{(1)} - \bar{x}_i}{\left[1 - \left(\frac{n}{n-1}\right)\left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right)'A^{-1}\left(\underset{\sim}{x}^{(1)}-\bar{\underset{\sim}{x}}\right)\right]^{\frac{1}{2}}}, \quad i = 1, 2, \ldots, p$$

then

$$1 - \frac{1}{n-1}\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\bar{x}}\right)'A^{-1}\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\bar{x}}\right) = \left[1 + \frac{1}{n-1-p}\,(a\underset{\sim}{Z})'A^{-1}(a\underset{\sim}{Z})\right]^{-1},$$

where

$$a\underset{\sim}{Z}' = \left(\sqrt{a_{11}}Z_1,\ \sqrt{a_{22}}Z_2,\ \ldots,\ \sqrt{a_{pp}}Z_{pp}\right).$$

Also,

$$\frac{\partial\left(\underset{\sim}{x}^{(1)}\right)}{\partial(\underset{\sim}{Z})} = \left(\frac{n-1}{n}\right)^{p/2} \frac{\prod_{i=1}^{p}\sqrt{a_{ii}}}{(n-p-1)^{p/2}} \left[1 + \frac{1}{n-p-1}\,(a\underset{\sim}{Z})'A^{-1}(a\underset{\sim}{Z})\right]^{-\frac{1}{2}(p+2)}$$

Equation (3.1.8) in $\underset{\sim}{Z}$ takes the form

$$f(Z_1,\ldots,Z_p)$$

$$= \frac{\Gamma\left(\frac{n-1}{2}\right)\prod_{i=1}^{p}\sqrt{a_{ii}}}{[\pi(n-1-p)]^{p/2}\Gamma\left(\frac{n-1-p}{2}\right)|A|^{\frac{1}{2}}} \left(1 + \frac{1}{n-1-p}\ \frac{\sum_{i}^{p}\sum_{j}^{p} A_{ij}\sqrt{a_{ii}a_{jj}}Z_iZ_j}{|A|}\right)^{-\frac{1}{2}(n-1)}$$

$$= \frac{\Gamma\left(\frac{n-1}{2}\right)}{[\pi(n-1-p)]^{p/2}\Gamma\left(\frac{n-1-p}{2}\right)|R|^{\frac{1}{2}}} \left(1 + \frac{1}{n-1-p}\,\underset{\sim}{Z}'R^{-1}\underset{\sim}{Z}\right)^{-\frac{1}{2}(n-1)}$$

where A_{ij} is the cofactor of a_{ij} and where R is the inverse of the sample correlation matrix R. As x_i's vary from $-\infty$ to h_i, the Z_i's vary from $-\infty$ to

$$Z_i(h_1,\ldots,h_p) = \left[\frac{n(n-1-p)}{n(n-1)a_{11}}\right]^{\frac{1}{2}} \frac{h_i-\bar{x}_i}{\left[1 - \frac{1}{n-1}\,(\underset{\sim}{h}-\underset{\sim}{\bar{x}})'A^{-1}(\underset{\sim}{h}-\underset{\sim}{\bar{x}})\right]^{\frac{1}{2}}},$$

Integral (3.1.10) is thus obtained.

VITA

Hong-i Yang

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE BEST ESTIMATE OF A MULTIVARIATE NORMAL DISTRIBUTION

Major Field: Statistics

Biographical:

Personal Data: Born in Mien-Yang, Hupeh, China, April 18, 1934, the son of Mr. and Mrs. Pi-shiu Yang.

Education: Attended elementary school in Hunan, China; attended junior high school in Hupeh, China; graduated from Changhua Senior High School, Taiwan, China, in 1952; received the Bachelor of Arts degree from the National Taiwan University, Taipei, Taiwan, in 1956; received the Bachelor of Arts degree from Montana State University at Missoula, Montana, in 1961; received the Master of Science degree at Ames, Iowa, in 1964; completed the requirements for the Doctor of Philosophy degree in July, 1974 at Oklahoma State University.

Professional Experience: Second Lieutenant in the Nationalist Chinese Army in Taiwan from August, 1956 to August, 1958; instructor and assistant professor of mathematics at Indiana Institute of Technology at Fort Wayne, Indiana, from 1964 to 1970; graduate assistant in the Department of Statistics, Oklahoma State University, Stillwater, Okahoma, from 1970 to 1974.